



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 173140

TO: Sheridan Swope
Location: REM-2B71/3C70
Art Unit: 1656
Thursday, December 08, 2005

Case Serial Number: 10/803530

From: Alex Waclawiw
Location: Biotech-Chem Library
Rem 1A71
Phone: 272-2534

Alexandra.waclawiw@uspto.gov

Search Notes

THIS PAGE BLANK (USPTO)

173140

STIC-Biotech/ChemLib

From: Swope, Sheridan
Sent: Friday, December 02, 2005 12:03 AM
To: STIC-Biotech/ChemLib
Subject: 10/803530

For 10/803530, pls search and interference search:

SID 2 against the NT and AA data bases.

Sheridan Swope, Ph.D.
Patent Examiner, AU 1656
Recombinant Enzymes
571-272-0943 (voice)
E02B71 Remsen Bld (Office)
E03C70 Remsen Bld (Mailbox)

Not MacLaurin

Point of Contact:
Alexandra Wacławiw
Technical Info. Specialist
CM 6A02 Tel 308-448

Searcher: _____
Searcher Phone: _____
Date Searcher Picked up: 12-5
Date completed: 12-8
Searcher Prep Time: 8
Online Time: 9

Type of Search
NA# _____ AA# 2
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: /
WWW/Internet: _____
Other (Specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 6, 2005, 20:10:51 ; Search time 6338 Seconds
(without alignments)
3794.862 Million cell updates/sec

Title: US-10-803-530-2
Perfect score: 2342
Sequence: 1 MDPDSDDPLNSLDVKPLRPK VYTKVSAYLNWYNNWKAEL 435

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 79147668 seqs, 27645789525 residues
Total number of hits satisfying chosen parameters: 158295336

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV=xlp
-MODEL=frame+p2n.model -DB=USPTO spooll_p/US10803530/runat_05122005_083217_2181/app_query.fasta_1.583
-Q=/cgn2_1/USPTO spooll_p/US10803530/runat_05122005_083217_2181/app_query.fasta_1.583
-DB=Pending Patents NA Main -QFMT=fastap -SUFFIX=rnpsm -MINMATCH=0.1 -LOOPCL=0
-LOOPTXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptp -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10803530@cgn2_1.11077@runat_05122005_083217_2181 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -XGAPOP=10 -YGAPEXT=0.5 -DELOP=5 -DELEXT=7

Database : Pending Patents NA Main:
1: /cgn2_6/ptodata/1/pna/PCTUSA COMB.seq.*
2: /cgn2_6/ptodata/1/pna/PCTUSB COMB.seq.*
3: /cgn2_6/ptodata/1/pna/PCTUSC COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US06 COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US075 COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US076 COMB.seq.*
7: /cgn2_6/ptodata/1/pna/US077 COMB.seq.*
8: /cgn2_6/ptodata/1/pna/US078 COMB.seq.*
9: /cgn2_6/ptodata/1/pna/US079 COMB.seq.*
10: /cgn2_6/ptodata/1/pna/US080 COMB.seq.*
11: /cgn2_6/ptodata/1/pna/US081 COMB.seq.*
12: /cgn2_6/ptodata/1/pna/US082 COMB.seq.*
13: /cgn2_6/ptodata/1/pna/US083 COMB.seq.*
14: /cgn2_6/ptodata/1/pna/US084 COMB.seq.*
15: /cgn2_6/ptodata/1/pna/US085 COMB.seq.*
16: /cgn2_6/ptodata/1/pna/US086 COMB.seq.*
17: /cgn2_6/ptodata/1/pna/US087 COMB.seq.*
18: /cgn2_6/ptodata/1/pna/US088 COMB.seq.*
19: /cgn2_6/ptodata/1/pna/US089 COMB.seq.*
20: /cgn2_6/ptodata/1/pna/US090 COMB.seq.*
21: /cgn2_6/ptodata/1/pna/US091 COMB.seq.*
22: /cgn2_6/ptodata/1/pna/US092 COMB.seq.*
23: /cgn2_6/ptodata/1/pna/US093 COMB.seq.*
24: /cgn2_6/ptodata/1/pna/US094 COMB.seq.*
25: /cgn2_6/ptodata/1/pna/US095A COMB.seq.*
26: /cgn2_6/ptodata/1/pna/US095B COMB.seq.*
27: /cgn2_6/ptodata/1/pna/US095C COMB.seq.*
28: /cgn2_6/ptodata/1/pna/US096A COMB.seq.*

29: /cgn2_6/ptodata/1/pna/US096B COMB.seq.*
30: /cgn2_6/ptodata/1/pna/US096C COMB.seq.*
31: /cgn2_6/ptodata/1/pna/US097A COMB.seq.*
32: /cgn2_6/ptodata/1/pna/US097B COMB.seq.*
33: /cgn2_6/ptodata/1/pna/US098A COMB.seq.*
34: /cgn2_6/ptodata/1/pna/US098B COMB.seq.*
35: /cgn2_6/ptodata/1/pna/US099A COMB.seq.*
36: /cgn2_6/ptodata/1/pna/US099B COMB.seq.*
37: /cgn2_6/ptodata/1/pna/US099C COMB.seq.*
38: /cgn2_6/ptodata/1/pna/US099D COMB.seq.*
39: /cgn2_6/ptodata/1/pna/US099E COMB.seq.*
40: /cgn2_6/ptodata/1/pna/US100A COMB.seq.*
41: /cgn2_6/ptodata/1/pna/US100B COMB.seq.*
42: /cgn2_6/ptodata/1/pna/US101 COMB.seq.*
43: /cgn2_6/ptodata/1/pna/US102A COMB.seq.*
44: /cgn2_6/ptodata/1/pna/US102B COMB.seq.*
45: /cgn2_6/ptodata/1/pna/US103A COMB.seq.*
46: /cgn2_6/ptodata/1/pna/US103B COMB.seq.*
47: /cgn2_6/ptodata/1/pna/US103C COMB.seq.*
48: /cgn2_6/ptodata/1/pna/US103D COMB.seq.*
49: /cgn2_6/ptodata/1/pna/US103E COMB.seq.*
50: /cgn2_6/ptodata/1/pna/US103F COMB.seq.*
51: /cgn2_6/ptodata/1/pna/US104 COMB.seq.*
52: /cgn2_6/ptodata/1/pna/US105 COMB.seq.*
53: /cgn2_6/ptodata/1/pna/US106A COMB.seq.*
54: /cgn2_6/ptodata/1/pna/US106B COMB.seq.*
55: /cgn2_6/ptodata/1/pna/US107A COMB.seq.*
56: /cgn2_6/ptodata/1/pna/US107B COMB.seq.*
57: /cgn2_6/ptodata/1/pna/US107C COMB.seq.*
58: /cgn2_6/ptodata/1/pna/US107D COMB.seq.*
59: /cgn2_6/ptodata/1/pna/US107E COMB.seq.*
60: /cgn2_6/ptodata/1/pna/US107F COMB.seq.*
61: /cgn2_6/ptodata/1/pna/US107G COMB.seq.*
62: /cgn2_6/ptodata/1/pna/US108 COMB.seq.*
63: /cgn2_6/ptodata/1/pna/US109A COMB.seq.*
64: /cgn2_6/ptodata/1/pna/US109B COMB.seq.*
65: /cgn2_6/ptodata/1/pna/US109C COMB.seq.*
66: /cgn2_6/ptodata/1/pna/US110A COMB.seq.*
67: /cgn2_6/ptodata/1/pna/US110B COMB.seq.*
68: /cgn2_6/ptodata/1/pna/US110C COMB.seq.*
69: /cgn2_6/ptodata/1/pna/US110D COMB.seq.*
70: /cgn2_6/ptodata/1/pna/US111A COMB.seq.*
71: /cgn2_6/ptodata/1/pna/US111B COMB.seq.*
72: /cgn2_6/ptodata/1/pna/US112 COMB.seq.*
73: /cgn2_6/ptodata/1/pna/US117 COMB.seq.*
74: /cgn2_6/ptodata/1/pna/US600 COMB.seq.*
75: /cgn2_6/ptodata/1/pna/US601 COMB.seq.*
76: /cgn2_6/ptodata/1/pna/US602A COMB.seq.*
77: /cgn2_6/ptodata/1/pna/US602B COMB.seq.*
78: /cgn2_6/ptodata/1/pna/US603 COMB.seq.*
79: /cgn2_6/ptodata/1/pna/US604A COMB.seq.*
80: /cgn2_6/ptodata/1/pna/US604B COMB.seq.*
81: /cgn2_6/ptodata/1/pna/US605 COMB.seq.*
82: /cgn2_6/ptodata/1/pna/US606 COMB.seq.*
83: /cgn2_6/ptodata/1/pna/US607 COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	2342	100.0	1305 40	US-10-030-688-1
2	2342	100.0	2154 28	US-09-606-680-3113
3	2338	99.8	2038 29	US-09-659-151-18
4	2338	99.8	2038 42	US-10-180-719-18
5	2338	99.8	2038 66	US-11-045-577-18
6	2338	99.8	2038 71	US-11-183-914-18
7	2337	99.8	1314 41	PCT-US02-19237-88
8	2337	99.8	1314 42	US-10-126-052A-448

Sequence 1, Appl
Sequence 3113, Ap
Sequence 18, Appl
Sequence 18, Appl
Sequence 18, Appl
Sequence 18, Appl
Sequence 88, Appl
Sequence 448, Appl

```
9 2337 99.8 1314 42 US-10-173-999-88 Sequence 88, Appl
10 2337 99.8 1314 44 US-10-295-027-133 Sequence 133, App
11 2337 99.8 1314 44 US-10-295-027-778 Sequence 778, App
12 2337 99.8 1314 44 US-10-295-027-790 Sequence 790, App
13 2337 99.8 1314 44 US-10-295-027-830 Sequence 830, App
14 2337 99.8 1314 44 US-10-295-027-979 Sequence 979, App
15 2337 99.8 2079 82 US-60-625-561-448 Sequence 448, App
16 2337 99.8 2104 3 PCT-US04-21227-1 Sequence 1, Appli
17 2337 99.8 2104 64 US-10-956-157-2292 Sequence 2, Appli
18 2337 99.8 2104 65 US-10-991-287-2 Sequence 2, Appli
19 2337 99.8 2104 65 US-10-991-287-2 Sequence 2, Appli
20 2337 99.8 2104 81 US-60-507-511-2292 Sequence 2, Appli
21 2337 99.8 2307 1 PCT-US02-07826-317 Sequence 2292, Ap
22 2337 99.8 2307 1 PCT-US02-07826-317 Sequence 217, App
23 2337 99.8 2307 40 US-10-097-340-317 Sequence 317, App
24 2337 99.8 2307 42 US-10-171-311-217 Sequence 317, App
25 2337 99.8 2307 66 US-11-050-926-317 Sequence 317, App
26 2337 99.8 2307 66 US-11-050-926-317 Sequence 317, App
27 2335 99.7 2121 1 PCT-US01-18568-1 Sequence 1, Appli
28 2335 99.7 2121 1 PCT-US01-18568-1 Sequence 1, Appli
29 2335 99.7 2121 62 US-10-803-530-1 Sequence 1, Appli
30 2333 99.6 2590 51 US-10-417-375-139 Sequence 139, App
31 2333 99.6 2590 51 US-10-417-375-139 Sequence 139, App
32 2333 99.6 2590 51 US-10-417-375A-139 Sequence 139, App
33 2333 99.6 2590 82 US-60-625-561-447 Sequence 447, App
34 2329 99.4 1479 40 US-10-030-688-3 Sequence 3, Appli
35 2329 99.4 2165 1 PCT-US02-08456-634 Sequence 634, App
36 2329 99.4 2165 42 US-10-101-510-634 Sequence 634, App
37 2329 99.4 2165 71 US-11-146-198-634 Sequence 634, App
38 2328 99.4 2108 3 PCT-US05-22501-4335 Sequence 4335, Ap
39 2328 99.4 2112 42 US-10-170-235-14349 Sequence 14349, A
40 2328 99.4 2627 51 US-10-417-375-141 Sequence 141, App
41 2328 99.4 2627 51 US-10-417-375A-141 Sequence 141, App
42 2328 99.4 2627 51 US-10-417-375B-141 Sequence 141, App
43 2328 99.4 2627 82 US-60-625-561-446 Sequence 446, App
44 2324 99.2 2070 3 PCT-US99-19655-2 Sequence 2, Appli
45 2324 99.2 2070 31 US-09-763-978A-2 Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-10-030-688-1
; Sequence 1, Application US/10030688
; GENERAL INFORMATION:
; APPLICANT: Merck Patent GmbH
; TITLE OF INVENTION: Seripancrin
; FILE REFERENCE: SeripancrinUHS
; CURRENT APPLICATION NUMBER: US/10/030,688
; CURRENT FILING DATE: 2002-01-14
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1305
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1305)
US-10-030-688-1
Alignment Scores:
Pred. No.: 0
Score: 2342.00 Length: 1305
Percent Similarity: 100.00% Matches: 435
Best Local Similarity: 100.00% Conservative: 0
Query Match: 100.00% Indels: 0
DB: 40 Gaps: 0

US-10-803-530-2 (1-435) x US-10-030-688-1 (1-1305)

Qy 1 MetAspProAspSerGlnProLeuAsnSerLeuAspValLeuProLeuArgLysPro 20
|||||
```

Qy	381	CysGlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGly	400
Db	1141	TGCAGGGTGACAGTGGTGGGCCCTCATGTACCAATCTACCAGTGGCATGTGGTGGC	1200
Qy	401	IleValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysVal	420
Db	1201	ATGTTAGCTGGGGCTATGGCTGGGGGGGCCGAGCACCCACGAGGATATACACCAAGTC	1260
Qy	421	SerAlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu	435
Db	1261	TCAGCCATTCTCAACTGGATCTCAATGTCTTGGAGGCTGAGCTG	1305

RESULT 2

```

US-09-606-680-3113
;
; Sequence 3113, Application US/09606680
; GENERAL INFORMATION:
; APPLICANT: Lloyd, Clare M.
; APPLICANT: Williamson, Mark
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.1131-001
; CURRENT APPLICATION NUMBER: US/09/606,680
; PRIORITY FILING DATE: 2000-06-27
; PRIORITY APPLICATION NUMBER: 60/141,227
; PRIOR FILING DATE: 1999-06-29
; PRIORITY APPLICATION NUMBER: 60/141,226
; PRIOR FILING DATE: 1999-06-29
; NUMBER OF SEQ ID NOS: 4394
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3113
; LENGTH: 2154
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-606-680-3113

```

; TITLE OF INVENTION: HUMAN PROTEASE MOLECULES
 ; NUMBER OF SEQUENCES: 24
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Dr.
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304
 ;
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/659,151
 ; FILING DATE: 11-Sep-2000
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 09/008,271
 ; FILING DATE: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Mohan-Peterson, Sheela
 ; REGISTRATION NUMBER: 41,201
 ; REFERENCE/DOCKET NUMBER: PF-0458 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 650-855-0555
 ; TELEFAX: 650-845-4166
 ; INFORMATION FOR SEQ ID NO: 18:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2038 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: COLNOT13
 ; CLONE: 1337018
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 18:
 ; US-09-659-151-18

Alignment Scores:
 Pred. No.: 0 Length: 2038
 Score: 2338.00 Matches: 434
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.77% Mismatches: 0
 Query Match: 99.83% Indels: 0
 DB: 29 Gaps: 0

US-10-803-530-2 (1-435) x US-09-659-151-18 (1-2038)

Qy 1 MetAspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysPro 20
 Db 200 ATGGATCCTGACAGTGATCAACCTCTGNACAGCCTCGATGTCAAACCCCTGGCGAAACCC 259
 Qy 21 ArgIleProMetGluThrPheArgLysValGlyIleProIleIleAlaLeuLeuSer 40
 Db 260 CGTATCCCATGTGGAGACCTTCAGAAAGTGGGGATCCCATCATCATAGCACTACTGAGC 319
 Qy 41 LeuAlaSerIleIleValValValValValValValValValValValValValVal 60
 Db 320 CTGGCGAGTATCATCTTGTTGTTCTCATCAAGGTGATTCTGGATAAATACTACTTC 379
 Qy 61 LeuCysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAsp 80
 Db 380 CTCTGGCGGAGCGCTCTCCACTTCATCCGAGGAGCAGCTGTGTGACGGAGAGCTGGAC 439
 Qy 81 CysProLeuGlyGluAspGluGluHisCysValLysSerPheProGluGlyProAlaVal 100
 Db 440 TGTCCTTGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTCCCGAAGGCGCTGCAGTG 499
 Qy 101 AlaValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsn 120
 Db 500 GCAGTCGCGCTCTCCAAGGACCGATCCACACTCAGGTGCTGAGCTCGGCCACAGGGAAC 559

Qy 121 TrpPheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGln 140
 Db 560 TGGTTCCTCTGCTGCTTTGACAACTTCACAGAGCTCTCGGTGAGACAGCCTGTAGGCAG 619
 Qy 141 MetGlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeu 160
 Db 620 ATGGGCTACAGCAGCAAAACCCACTTTTCTAGAGCTGTGGAGATTGGCCCGAGACCATG 679
 Qy 161 AspValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyPro 180
 Db 680 GATGTTGTTGAATTCACAGAAACACAGCAGGAGCTTCGCATGCGGAACCTCAAGTGGGCC 739
 Qy 181 CysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThr 200
 Db 740 TGTCTCTCAGGCTCCCTGGTCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 799
 Qy 201 ProArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIle 220
 Db 800 CCCGCTGTGGTGGGAGGAGGCTCTGTGATTTCTTGGCTTGGCAGGTGAGCATC 859
 Qy 221 GlnTyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThr 240
 Db 860 CAGTACGACAAACAGCAGCAGCTCTGTGGAGGAGCATCTGGACCCCACTGGGTCTTCACG 919
 Qy 241 AlaAlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySer 260
 Db 920 GCAGCCCACTGCTTCAGGAACATACCGATGTTTCACTGGGAAGTGGCGGAGGCTCA 979
 Qy 261 AspLysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleIleGluPheAsn 280
 Db 980 GACAAACTGGGAGCTTCCCATCCCTGGCTGTGGCCAGATCATCATCATTTGATTTCAAC 1039
 Qy 281 ProMetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPhe 300
 Db 1040 CCCATGTACCCCAAGACATGACATCGCCCTCATGAAGCTGAGTTCCTCCACTCTTC 1099
 Qy 301 SerGlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThr 320
 Db 1100 TCAGGCACAGTCAGGCCCATCTGTGCTGCTTCTTTGATGAGGAGCTCACTCCAGCCACC 1159
 Qy 321 ProLeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyLysMetSerAspIle 340
 Db 1160 CCACCTCTGATCATTTGGATGGGCTTTACGAAGCAGAAATGGAGGAGAGATGTCACATA 1219
 Qy 341 LeuLeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaTyr 360
 Db 1220 CTGCTGCAGGCTCAGTCCAGGTCAATTGACAGCACACGCTGCAATGSCAGACGATCGTAC 1279
 Qy 361 GlnGlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThr 380
 Db 1280 CAGGGGGAAGTCACCGAGAAGATGATGTGTGACGCATCCCGGAAGGGGGTGTGACACC 1339
 Qy 381 CysGlnGlyAspSerGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGly 400
 Db 1340 TGCCAGGTGACAGTGGTGGGCCCCCTGATGTACCAATCTGACCACTGTCATGTGGTGGGC 1399
 Qy 401 IleValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrIleVal 420
 Db 1400 ATCGTTAGCTGGGCTATGGCTGGGGGGCCCCAGACCCCGAGGATATACCAAGGTC 1459
 Qy 421 SerAlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
 Db 1460 TCAGCCTATCTCAACTGATCTACAATGTCTGGAAGGCTGAGCTG 1504

RESULT 4

; US-10-180-719-18
 ; Sequence 18, Application US/10180719
 ; GENERAL INFORMATION:
 ; APPLICANT: Bandman, Olga
 ; Hillman, Jennifer L.
 ; Yue, Henry
 ; Guegler, Karl J.
 ; Corley, Neil C.

; Tang, Tom Y.
 ; Shah, Purvi
 ; TITLE OF INVENTION: HUMAN PROTEASE MOLECULES
 ; NUMBER OF SEQUENCES: 24
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Dr.
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/10/180,719
 ; FILING DATE: 25-Jun-2002
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/008,271
 ; FILING DATE: 16-Jan-1998
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Mohan-Peterson, Sheela
 ; REGISTRATION NUMBER: 41,201
 ; REFERENCE/DOCKET NUMBER: PF-0458 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 650-855-0555
 ; TELEFAX: 650-845-4166
 ; INFORMATION FOR SEQ ID NO: 18:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2038 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: COLNNOT13
 ; CLONE: 1337018
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 18:
 ; US-10-180-719-18

Alignment Scores:
 Pred. No.: 0 Length: 2038
 Score: 2338.00 Matches: 434
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.77% Mismatches: 0
 Query Match: 99.83% Indels: 0
 DB: 42 Gaps: 0

US-10-803-530-2 (1-435) x US-10-180-719-18 (1-2038)

QY 1 MetAspProAspSerAspGlnProLeuAenSerLeuAspValLysProLeuArgLysPro 20
 DB 200 ATGGATCCTGACAGTATCAACCTCTGAACAGCTCGATGTCAACCCCTGGCGAAACCC 259
 QY 21 ArgIleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSer 40
 DB 260 CGTATCCCCATGAGACCTTCAGAAAGTGGGATCCCCATCATCATAGCACTACTGAGC 319
 QY 41 LeuAlaSerIleIleValValValValLeuIleLysValIleLeuAspLysTyrPhe 60
 DB 320 CTGGCAGTATCATATTGGTGTCTCATCAAGTGATTCGGATAAATACTACTTC 379
 QY 61 LeuCysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAsp 80
 DB 380 CTCCTGGCGGAGCCTCTCCACTTCATCCCGAGAGCAGCTGTGTACCGAGAGCTGGAC 439
 QY 81 CysProLeuGlyGluAspGluHisCysValLysSerPheProGluGlyProAlaVal 100
 DB 440 TGTCCCTTGGGGAGACGAGGAGCACTGTGTCAAGAGCTTCCCGAGGGCTCGAGTG 499
 QY 101 AlaValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsn 120

DB 500 GCAGTCCGCTCTCCNAGGACCGATCCACACTGCAGGTGCTGGACTCGGCCACAGGGAAC 559
 QY 121 TrpPheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGln 140
 DB 560 TGGTCTCTGCTGCTTTTCGACAACTTCACAGAACTCTCGCTGAGACAGCTGTAGGCAG 619
 QY 141 MetGlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeu 160
 DB 620 ATGGGCTACAGAGCAAAACCCACTTTTCAGAGCTGTGGAGATTGGCCCAAGACAGGATCTG 679
 QY 161 AspValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyPro 180
 DB 680 GATGTTGTTGAATTCACAGAAAACAGCCAGAGCTTCGCATGCGGAACTCAAGTGGGCC 739
 QY 181 CysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThr 200
 DB 740 TGTCTCTCAGGCTCCCTGGTCTCCCTGACACTGTCTTGTGCTGTGGGAGAGAGCCCTGAAGAC 799
 QY 201 ProArgValValGlyGlyGluAlaSerValAspSerTrpProTrpGlnValSerIle 220
 DB 800 CCCCCTGTGGTGGGTGGGAGAGCCCTCTGTGGATTCTTGGCCCTTGGCAGGTCAAGCATC 859
 QY 221 GlnTyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThr 240
 DB 860 CAGTACGACAAACAGCAGCAGTCTGTGGAGGAGCATCTTGGACCCCACTGGGTCTCTCAG 919
 QY 241 AlaAlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValAlaArgAlaGlySer 260
 DB 920 GCAGCCCCACTGCTTCAGGAAACATACCGATGTGTTCAACTGGAGAGTGGCGGAGCTCA 979
 QY 261 AspLysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleIleGluPheAsn 280
 DB 980 GACAAACTGGGAGCTTCCCATCCCTGCTGTGGCCCAAGATCATCATTAATTCATCAAC 1039
 QY 281 ProMetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPhe 300
 DB 1040 CCCATGTACCCCAAGACAGCATCGCCCTCATGAAGCTGCAGTTCACACTCACTTTC 1099
 QY 301 SerGlyThrValArgProIleCysLeuProPheAspGluGluLeuThrProAlaThr 320
 DB 1100 TCAGGCACAGTCAGGCCCATCTGTCTGCCCTTCTTTGATGAGGAGCTCACTCCAGCCACC 1159
 QY 321 ProLeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIle 340
 DB 1160 CCACTCTGGATCATTTGGATGGGCTTTACGAAAGAGAAATGGAGGAAAGATGTCTGACATA 1219
 QY 341 LeuLeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyr 360
 DB 1220 CTGCTGCAAGCGCTCAGTTCAGGTCAATTGACAGCACACGGTGCATGACAGACGATGCGTAC 1279
 QY 361 GlnGlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThr 380
 DB 1280 CAGGGGGAAGTCAACGAGAAGATGATGTGTGAGGATCCCGAAGGGGGTGTGGACACC 1339
 QY 381 CysGlnGlyAspSerGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGly 400
 DB 1340 TGCCAGGGTGACAGTGGTGGGCCCTGTGTACCAATCTGACCACTGACCACTGGCATGTGGGG 1399
 QY 401 IleValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValThrThrLysVal 420
 DB 1400 ATCGTGTAGTGGGGCTATGGCTGCGGGGGGCCAGCAGCCCCAGGAGATATACACCAAGGTC 1459
 QY 421 SerAlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
 DB 1460 TCAGCTATCTCACTGGATCTACAAATGTCTGGAGGCTGAGCTG 1504

RESULT 5

US-11-045-577-18
 ; Sequence 18, Application US/11045577
 ; GENERAL INFORMATION:
 ; APPLICANT: Bandman, Olga
 ; Hillman, Jennifer L.
 ; Yue, Henry

;; Guegler, Karl J.
;; Corley, Neil C.
;; Tang, Tom Y.
;; Shah, Purvi
;; TITLE OF INVENTION: HUMAN PROTEASE MOLECULES
;; NUMBER OF SEQUENCES: 24
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Incyte Pharmaceuticals, Inc.
;; STREET: 3174 Porter Dr.
;; CITY: Palo Alto
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 94304
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: DOS
;; SOFTWARE: FastSeq for Windows Version 2.0
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/11/045,577
;; FILING DATE: 27-Jan-2005
;; PRIORITY APPLICATION DATA:
;; APPLICATION NUMBER: US/09/659,151
;; FILING DATE: 11-Sep-2000
;; APPLICATION NUMBER: 09/008,271
;; FILING DATE: <Unknown>
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Mohan-Peterson, Sheela
;; REGISTRATION NUMBER: 41,201
;; REFERENCE/DOCKET NUMBER: PF-0458 US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 650-855-0555
;; TELEFAX: 650-845-4166
;; INFORMATION FOR SEQ ID NO: 18:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2038 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; IMMEDIATE SOURCE:
;; LIBRARY: COLNNOT13
;; CLONE: 1337018
;; SEQUENCE DESCRIPTION: SEQ ID NO: 18:

US-11-045-577-18

Alignment Scores:
Pred. No.: 0 Length: 2038
Score: 2338.00 Matches: 434
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.77% Mismatches: 0
Query Match: 99.83% Indels: 0
DB: 66 Gaps: 0

US-10-803-530-2 (1-435) x US-11-045-577-18 (1-2038)

Qy 1 MetAspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysPro 20
Db 200 ATGGATCTGACATGATCAACCTCTGAACAGCTTCGATGTAACCCCTCGCGAAACCC 259
Qy 21 ArgLeuProMetGluThrPheArgLysValGlyLeuProLeuLeuLeuLeuLeuLeu 40
Db 260 CGTATCCCATGGAGACCTTCAGAAAGGTGGGATCCCCATCATAGCACTACTGAGC 319
Qy 41 LeuAlaSerIleIleValValValValValValValValValValValValValVal 60
Db 320 CTGCGGAGTATCATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 379
Qy 61 LeuCysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAsp 80
Db 380 CTCTCGGGCAGCTCTCCACTTCATCCCGAGGAGCAGCTGTGACGGAGAGCTGGAC 439
Qy 81 CysProLeuGlyGluAspGluGluHisCysValLysSerPheProGluGlyProAlaVal 100

RESULT 6

US-11-183-914-18

; Sequence 18, Application US/11183914

Db 440 TGTCCCTTGGGGGAGGACGAGGACACTGTGTCTCAAGAGCTTCCCGAAGGCGCTGCAGTG 499
Qy 101 AlaValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsn 120
Db 500 GCAGTCCGCTCTCCAAAGGACCGATCCACACTGCAGGTGTGGACTGGCCACAGGNAAC 559
Qy 121 TrpPheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGln 140
Db 560 TGGTCTCTGCTCTGTTTTCGACAACTTCACAGAACTCTCGCTGAGACAGCCTGTAGGCAG 619
Qy 141 MetGlyTyrSerSerLysProThrPheArgAlaValGluLeuGlyProAspGlnAspLeu 160
Db 620 ATGGGCTACAGCAGCAAAACCCACTTTCAGAGCTGTGGAGATTGGCCAGACCCAGATCTG 679
Qy 161 AspValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyPro 180
Db 680 GATGTTGTTGAAATTCACAGAAACAGCCAGGAGCTTCGTCATGCGAATCAAGTGGGCC 739
Qy 181 CysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThr 200
Db 740 TGTCTCTCAGGCTCCCTGGTCTCCCTGCACTGTCTTGGCTTGGCGAGAGCCTGAAGACC 799
Qy 201 ProArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIle 220
Db 800 CCCCCTGTGTGGTGGGAGGAGGCTCTGTGGAGGAGCATCTGGACCCCACTGGGCTCTCACG 859
Qy 221 GlnTyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThr 240
Db 860 CAGTACGACAAACAGCAGCGCTCTGTGGAGGAGCATCTGGACCCCACTGGGCTCTCACG 919
Qy 241 AlaAlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySer 260
Db 920 GCAGCCCACTGCTTCAGGAAACATACCGATGTGTTCACTTGAAGTGGCGGAGGCTCA 979
Qy 261 AspLysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsn 280
Db 980 GACAAACTGGGAGCTTCCCATCCCTGGTGTGGCCCAAGATCATCATATTGAATTCAC 1039
Qy 281 ProMetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPhe 300
Db 1040 CCAATGTACCCAAAGACAATGACATCGCCCTCATGAAGCTGCACTGCCACTCTCTTC 1099
Qy 301 SerGlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThr 320
Db 1100 TCAGGCACAGTCAGGCCCATCTGTCTGCCCTTCTTGTAGAGAGCTCACTCCAGCCACC 1159
Qy 321 ProLeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIle 340
Db 1160 CCACCTCTGGATCATTTGGATGGGCTTTAGAAAGCAATGGAGGAAGATGTCTGACATA 1219
Qy 341 LeuLeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyr 360
Db 1220 CTGCTCAGGCGCTCAGTCCAGGTCAATTGACACACACCGTGCATGCGAGACATGCGTAC 1279
Qy 361 GlnGlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThr 380
Db 1280 CAGGGGGAAGTCACCGAGAAGATGATGTGTGCAGCATCCCGGAAGGGGTGTGGACACC 1339
Qy 381 CysGlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGly 400
Db 1340 TGCCAGGGTGACAGTGGTGGGCCCCCTGATGTACCAATCTGACCACTGGCATGTGGTGG 1399
Qy 401 IleValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysVal 420
Db 1400 ATCGTTAGCTGGGCTATGGCTGGGGGGCCCGAGCACCCAGAGAGTATACCAAGGTTC 1459
Qy 421 SerAlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db 1460 TCAGCCTATCTCACTGGATCTACAATGTCTGGAAGGCTGAGCTG 1504

GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Hillman, Jennifer L.
APPLICANT: Yue, Henry
APPLICANT: Guegler, Karl J.
APPLICANT: Corley, Neil C.
APPLICANT: Tang, Tom Y.
APPLICANT: Shah, Purvi
TITLE OF INVENTION: HUMAN PROTEASE MOLECULES
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/11/183,914
FILING DATE: 19-JULY-2005
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/008,271
FILING DATE: 16-Jan-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Mohan-Peterson, Sheela
REGISTRATION NUMBER: 41,201
REFERENCE/DOCKET NUMBER: PF-0458 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 2038 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: COLNOT13
CLONE: 1337018
US-11-183-914-18

Alignment Scores:
Pred. No.: 0 Length: 2038
Score: 2338.00 Matches: 434
Percent Similarity: 100.00% Conservatives: 1
Best Local Similarity: 99.77% Mismatches: 0
Query Match: 99.83% Indels: 0
DB: 71 Gaps: 0

US-10-803-530-2 (1-435) x US-11-183-914-18 (1-2038)

QY 1 MetAppProAspSerGlnProLeuAnSerLeuAspValLysProLeuArgLysPro 20
DB 200 ATGGATCCCTGACAGTGCATCAACCTCTGAAACAGCTCGATGTCAACCCCTGGCAACCC 259

QY 21 ArgIleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuSer 40
DB 260 CGTATCCCCATGGAGACCTTCAGAAAGGTGGGATCCCATCATCATAGCTACTGAGC 319

QY 41 LeuAlaSerIleIleIleValValLeuIleValLysValIleLeuAspLysValTyrPhe 60
DB 320 CTGGCGAGTATCATCATTTGGTTGTCTCATCAAGGTGATCTGTGATTAATACTACTTC 379

QY 61 LeuCysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAsp 80
DB 1460 TCAGGCTATCTCAACTGGATCTACAATGTCTGGAAGGCTGAGCTG 1504

DB 380 CTCGGGGCAGCCTCTCCACTTCCATCCGAGGAGCAGCTGTGTGACGAGAGCTGGAC 439

QY 81 CysProLeuGlyGluAspGluGluHisCysValLysSerPheProGluGlyProAlaVal 100
DB 440 TGTCCCTTGGGGAGGAGCAGGAGCACTGTGTCAAGAGCTTCCCGAAGGGCTCGAGTG 499

QY 101 AlaValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAen 120
DB 500 GCAGTCCCGCCTCTCCAGGACCGATCCACACTGCAAGAGCTCTCGCTGACAGAGCTGTAGGCAG 559

QY 121 TrpPheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGln 140
DB 560 TGGTCTCTGCTGCTTCGACCACTTCACAGAGCTCTCGCTGACAGAGCTGTAGGCAG 619

QY 141 MetGlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeu 160
DB 620 ATGGGCTACAGCAGCAAAACCCACTTTCAGAGCTGTGGAGATTGGCCACAGACGATCTG 679

QY 161 AspValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyPro 180
DB 680 GATGTTGTGAAATCACAGAAACAGCCAGGAGCTTCGCATGCGGAACCTCAAGTGGGCC 739

QY 181 CysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThr 200
DB 740 TGTCTCTCAGGCTCCCTGGTCTCCTCTGCTGTCTGCTGTGGGAGAGCTGAAGACC 799

QY 201 ProArgValValGlyGlyGluAlaSerValAspSerTrpProTrpGlnValSerIle 220
DB 800 CCCCCTGTGGTGGGAGGAGGAGGCTCTGTGGATTCTTGGCCTTGGCAGGTGAGCATC 859

QY 221 GlnTyrAspLysGlnHisValCysGlySerIleLeuAspProHisTrpValLeuThr 240
DB 860 CAGTACGACAAACAGCACGCTCTGTGGAGGAGCATCTCTGGACCCCTCCCTGCTCAGC 919

QY 241 AlaAlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySer 260
DB 920 GCAGCCCACTGCTTCAGGAACATACCATGTGTCACTGGAAGGTGGGGCAGGCTCA 979

QY 261 AspLysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsn 280
DB 980 GACAACTGGGCAGCTTCCCATCCCTGCTGTGGCCAAGATCATCATTAATTCAC 1039

QY 281 ProMetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPhe 300
DB 1040 CCCATGTATCCCAACAAACATGACATCGCCCTCATGAAGCTGCAGTCTCCACTTC 1099

QY 301 SerGlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThr 320
DB 1100 TCAGGCACAGTCAGGCCCATCTGTCTGCCCTTCTTGTGATGAGGAGCTCACTCAGCCACC 1159

QY 321 ProLeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIle 340
DB 1160 CCACCTCTGATCATTTGGATGGGCTTTACGAGCAGATGGAGGAGAGATGTCTGACATA 1219

QY 341 LeuLeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaTyr 360
DB 1220 CTGCTGCAGGCGTCAGTCCAGGTCAATGACAGCACACGTCGCAATGACAGCATGCGTAC 1279

QY 361 GlnGlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyValLeuAspThr 380
DB 1280 CAGGGGGAAGTCACCGAGAAGATGATGTGTGAGGATCTCCCGAAGGGGTGTGGACACC 1339

QY 381 CysGlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGly 400
DB 1340 TGCCAGGTGACAGTGTGGGCCCTCTGATACCAATCTGACAGTGGCATGTGTGGTGGC 1399

QY 401 IleValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysVal 420
DB 1400 ATCGTTAGTGGGCTATGGCTGCGGGGCCCGAGCACCCAGGAGTATACACCAAGTTC 1459

QY 421 SerAlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
DB 1460 TCAGGCTATCTCAACTGGATCTACAATGTCTGGAAGGCTGAGCTG 1504

RESULT 7

PCT-US02-19297-88
; Sequence 88, Application PC/TUS0219297
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Eos Biotechnology Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Ovarian Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Ovarian
; FILE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002420PC
; CURRENT APPLICATION NUMBER: PCT/US02/19297
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US 60/299,234
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/315,287
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: US 60/317,544
; PRIOR FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 164
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 88
; LENGTH: 1314
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US02-19297-88

Alignment Scores:

Pred. No.:	0	Length:	1314
Score:	2337.00	Matches:	434
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	99.79%	Indels:	0
DB:	1	Gaps:	0

US-10-803-530-2 (1-435) x PCT-US02-19297-88 (1-1314)

QY	2	AspProAspSerAspGlnProLeuAsnSerLeuAspVallysProLeuArgLysProArg	21
Db	10	GATCTGACAGTATCAACCTCTGAACAGCCTCGATGTCAAAACCCCTGCGCAACCCCGT	69
QY	22	IleProMetGluThrPheArgLysValGlyIleProIleIleIleIleAlaLeuLeuSerLeu	41
Db	70	ATCCCCATGGAGACCTTCAGAAAGGTGGGATCCCCATCATAGCACTACTGAGCCTG	129
QY	42	AlaSerIleIleIleValValValLeuIleLysValIleLeuAspLysTyrTyrPheLeu	61
Db	130	CGAGATATCATGTGTGTGTCTCTCATCAAGGTGATCTGGATAAATACTACTTCTCTC	189
QY	62	CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys	81
Db	190	TGCGGGCAGCCTCTCCACTTCATCCGAGGAGCAGCTGTGTGACGGAGAGCTGGACTGT	249
QY	82	ProLeuGlyGluAspGluHisCysVallysSerPheProGluGlyProAlaValAla	101
Db	250	CCCTTGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTCCCCGAAGGGCTGCGAGTGGCA	309
QY	102	ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp	121
Db	310	GTCCGCTCTCCAAAGGACCATTCACACTCAGGTGCTGGACTCGGCCACAGGGAATGG	369
QY	122	PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet	141
Db	370	TTCTCTGCCTGTTTCGACAACTTCACAGAGCTCTCGCTGAGACAGCCTGTAGCGAGATG	429
QY	142	GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp	161
Db	430	GGCTACAGCAACCAACCTTTTCAGAGCTGTGGAGATTGGCCACAGGATCTGGAT	489

QY	162	ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys	181
Db	490	GTTGTTGAATCAAGAAAACAGCCAGAGCTTCGCATTCGGAACCTCAAGTGGGCGCTGT	549
QY	182	LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro	201
Db	550	CTCTCAGGCTCCCTGGTCTCCCTGCATGTCTTCCCTGTGGGAAGAGCCTGAAGACCCCC	609
QY	202	ArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleGln	221
Db	610	CGTGTGTGGTGGGAGGAGGCTCTGTGATCTTTGGCTTGGCAGGTCAGATCCAG	669
QY	222	TyrAspLysGlnHisValCysGlySerIleLeuAspProHisTrpValLeuThrAla	241
Db	670	TACGACAAACAGCACGCTCTGTGGAGGAGCATCTGGACCCCACTGGTGGTCTCACGGCA	729
QY	242	AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp	261
Db	730	GCCCACTGCTTCAGGAACATACCGATGTGTTCACTGGGAAGGTGCGGCGGCTCAGAC	789
QY	262	LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleIleGluPheAsnPro	281
Db	790	AAACTGGGAGCTTCCCATCCCTGGCTGTGCCAAGATCATCATTTGAATTCAACCCCC	849
QY	282	MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer	301
Db	850	ATGTACCCCAAGACAAATGACATCGCCCTCATGAAGCTGCAGTTCCTCCACTCTCTCA	909
QY	302	GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro	321
Db	910	GGCACAGTCAGGCCCATCTGTCTGCCCTTCTTTGATGAGGAGCTCACTCCAGCCACCCCA	969
QY	322	LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu	341
Db	970	CTCTGGATCATTTGGATGGGCTTTTACGAAGCAGATGGAGGAAGATGTCTGACATCTG	1029
QY	342	LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaTyrGln	361
Db	1030	CTGCAGGCTCAGTCCAGGTCAATTGACGACACGCGTCAATGACACCATGCGTACCAAG	1089
QY	362	GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys	381
Db	1090	GGGGAAGTCACGAGAAGATGATGTGTGAGGATCCCGAAGGGGGTGTGGACACTGC	1149
QY	382	GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle	401
Db	1150	CAGGGTGACAGTGGTGGGCTTGTATGATGATGATGATGATGATGATGATGATGATGATG	1209
QY	402	ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSer	421
Db	1210	GTTAGCTGGGGCTATGGCTGCGGGGGCCGAGACCCCAAGGAGTATACCAAGGTCTCA	1269
QY	422	AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu	435
Db	1270	GCCTATCTCACTGGATCTACAATGTCTGGAAGGCTGAGCTG	1311

RESULT 8

US-10-126-052A-448
; Sequence 448, Application US/10126052A
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Murray, Richard
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Lung Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Lung Cancer
; FILE OF INVENTION: Methods of Screening for Modulators of Lung Cancer
; FILE REFERENCE: 018501-001530US
; CURRENT APPLICATION NUMBER: US/10/126,052A
; CURRENT FILING DATE: 2002-04-18
; PRIOR APPLICATION NUMBER: US 60/284,770
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: US 60/290,492
; PRIOR FILING DATE: 2001-05-10

; PRIOR APPLICATION NUMBER: US 60/339,245
 ; PRIOR FILING DATE: 2001-11-09
 ; PRIOR APPLICATION NUMBER: US 60/350,666
 ; PRIOR FILING DATE: 2001-11-13
 ; PRIOR APPLICATION NUMBER: US 60/334,370
 ; PRIOR FILING DATE: 2001-11-29
 ; PRIOR APPLICATION NUMBER: US 60/372,246
 ; PRIOR FILING DATE: 2002-04-12
 ; NUMBER OF SEQ ID NOS: 691
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 448
 ; LENGTH: 1314
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-126-052A-448

Alignment Scores:
 Pred. No.: 0 Length: 1314
 Score: 2337.00 Matches: 434
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 99.79% Indels: 0
 DB: 42 Gaps: 0

US-10-803-530-2 (1-435) x US-10-126-052A-448 (1-1314)

QY 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysProArg 21
 DB 10 GATCTGACAGTGATCAACCTCTGAACAGCCTCGATGTCAAACCCCTCGCGAACCCTCGT 69
 QY 22 IleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSerLeu 41
 DB 70 ATCCCATGAGACCTTCAGAAAGGTGGGATCCCATCATCATATGACACTACTGAGCCTG 129
 QY 42 AlaSerIleIleIleValValValLeuIleLysValIleLeuAspLysTyrTyrPheLeu 61
 DB 130 GCGAGTATCATCATTTGTGTTTCTCTCATCAAGGTGATTTCTGGATAATACTACTTCTC 189
 QY 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyLeuLeuAspCys 81
 DB 190 TCGGGGACCTCTCCACTTCATCCCGAGAGAGCAGCTGTGTGACGAGAGCTGGACTGT 249
 QY 82 ProLeuGlyGluAspGlnGluHisCysValLysSerPheProGluGlyProAlaValAla 101
 DB 250 CCCTTTGGGGAGACGAGGAGCACTGTGTCAAGAGCTTCCCGAAGGGCTCGAGTGGCA 309
 QY 102 ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp 121
 DB 310 GTCCGCTCTCCAGGACCGATCCACACTGCGAGGTCTGGACTCGGCCACAGGGAACTGG 369
 QY 122 PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
 DB 370 TTCTCTGCTGTTTCGCAACTTCACAGAGCTCTGCTGACAGCCTGTAGGCAGATG 429
 QY 142 GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
 DB 430 GGCTACAGCAGCAACCCACTTTTCAGAGCTGTGGAGATTGGCCACAGCAGGATCTGGAT 489
 QY 162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
 DB 490 GTTGTGTAATCACAGAAACAGCCAGGAGCTTCGCATCGGAACCTCAAGTGGGCCCTGT 549
 QY 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro 201
 DB 550 CTCTAGGCTCCCTGCTCTCCCTGCACTGTCTTGGCTGTGGGAAGAGCTTGAAGACCCCC 609
 QY 202 ArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
 DB 610 CGTGTGTTGGTGGGAGGAGGCTCTGTGGATTCTTGGCTTGGCAGGTTCAGATCCAG 669
 QY 222 TyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAla 241
 DB 670 TACGACAAACAGCACGCTGTGTGGAGGAGCATCTTGGACCCCACTCTGGTCTCTCACGGCA 729

QY 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValAlaGlySerAsp 261
 DB 730 GCCCACTGCTTCAGGAAACATACCGATGTGTTCACCTGGAAGGTGCGGGCAGGCTCAGAC 789
 QY 262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleIleGluPheAsnPro 281
 DB 790 AAATCTGGGACGCTCCCATCCCTGGCTGTGGCCAGATCATCATCATTAATTCAACCCC 849
 QY 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
 DB 850 ATGTACCCCAAGACAATGACATGCCCTCATGAAGCTGCAGTTCCTCACTTCTCTCA 909
 QY 302 GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro 321
 DB 910 GGCACAGTCAGGCCCATCTGTCTGCCCTTCTTTGATGAGGAGCTCACTCCAGCCACCCA 969
 QY 322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu 341
 DB 970 CTCTGGATCATGGATGGGCTTTACGAAGCAGAAATGGAGGAGATGTCTGACATCTG 1029
 QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaTyrGln 361
 DB 1030 CTGCAGGCGTCAGTCCAGTCAATTGACAGCACACGGTGCATTCAGACGATCGTACCCAG 1089
 QY 362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys 381
 DB 1090 GGGGAAGTCACCGAGAAGATGATGTGCGAGCATCCCGAAGGGGGTGTGGACACCTGC 1149
 QY 382 GlnGlyAspSerGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle 401
 DB 1150 CAGGTGACAGTGTGGGCCCCCTGATGTACCAATCTGACCATGTGTCATGTGGGCACTC 1209
 QY 402 ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSer 421
 DB 1210 GTTAGCTGGGCTATGCTGCGGGGCGCCGAGCACCCAGGAGTATACCAAGGTCTCA 1269
 QY 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
 DB 1270 GCCTATCTCACTGGATCTACAATGTCTGGAAGGCTGAGCTG 1311

RESULT 9

; Sequence 88, Application US/10173999
 ; GENERAL INFORMATION:
 ; APPLICANT: Mack, David H.
 ; APPLICANT: Gish, Kurt C.
 ; APPLICANT: Eos Biotechnology, Inc.
 ; TITLE OF INVENTION: Methods of Diagnosis of Ovarian Cancer, Compositions
 ; TITLE OF INVENTION: and Methods of Screening for Modulators of Ovarian
 ; FILE REFERENCE: 018501-002420US
 ; CURRENT APPLICATION NUMBER: US/10/173,999
 ; CURRENT FILING DATE: 2002-06-17
 ; PRIOR APPLICATION NUMBER: US 60/299,234
 ; PRIOR FILING DATE: 2001-06-18
 ; PRIOR APPLICATION NUMBER: US 60/315,287
 ; PRIOR FILING DATE: 2001-08-27
 ; PRIOR APPLICATION NUMBER: US 60/350,666
 ; PRIOR FILING DATE: 2001-11-13
 ; PRIOR APPLICATION NUMBER: US 60/372,246
 ; PRIOR FILING DATE: 2001-04-12
 ; NUMBER OF SEQ ID NOS: 163
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 88
 ; LENGTH: 1314
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-173-999-88

Alignment Scores: 0 Length: 1314
 Pred. No.: 2337.00 Matches: 434
 Score:

Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 99.79%
DB: 42
Gaps: 0
Conservative: 0
Mismatch: 0
Indels: 0

US-10-803-530-2 (1-435) x US-10-173-999-88 (1-1314)

QY 2 AspProAspSerGlnProLeuAsnSerLeuAspValysProLeuArgLysProArg 21
DB 10 GATCCTTGACAGTATCAACCTCTGAACAGCCTCGATGTCAAAACCCCTCGGCAACCCCGT 69
QY 22 IleProMetGluThrPheArgLysValGlyIleProIleIleIleIleAlaLeuSerLeu 41
DB 70 ATCCCATGGAGACCTTCAGAAAGGTGGGATCCCATCATCATAGCACTACTAGACCTG 129
QY 42 AlaSerIleIleIleValValLeuIleLysValIleLysValIleLysValIleLysValIle 61
DB 130 GCGAGTATCATCTGCTGTGCTCATCAAGGTGATCTCGATTAATACTACTTCTCTC 189
QY 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyIleLeuAspCys 81
DB 190 TCGGGCAGCCCTTCACACTTCATCCGAGAGAGCAGCTGTGTGACGGAGAGCTGGACTGT 249
QY 82 ProLeuGlyGluAspGluHisCysValLysSerPheProGluGlyProAlaValAla 101
DB 250 CCTTGGGGAGAGCAGAGGACACTGTGTCAAGAGCTTCCCGAAGGGCCCTGCAAGTGGCA 309
QY 102 ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp 121
DB 310 GTCCGCTCTCCAAAGACCGATCCACACTGCAGGTGCTGGACTCGGCCACAGGAGACTGG 369
QY 122 PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
DB 370 TTCTCTGCTGTTTCGACAACTTCACAGAGCTCTCGCTGACAGAGCTGTAGGCAGATG 429
QY 142 GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
DB 430 GGCTACAGCAGCAAAACCCACTTCAGAGCTGTGGAGATGGCCCGAGAGAGCTGTGAT 489
QY 162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
DB 490 GTTGTGTAATCACAGAAACAGCCAGGAGCTTCGATCGCGAACTCAAGTGGGCCCTGT 549
QY 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro 201
DB 550 CTCTCAGGCTCCCTGGTCTCCCTGACACTGTCTGCTGTGGAGAGAGCCTGAAGACCCCC 609
QY 202 ArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
DB 610 CGTGTGTGGTGGGAGAGGCTCTGTGATCTTGGCTTGGCAGGTTCAGATCCAG 669
QY 222 TyrAspLysGlnHisValCysGlySerIleLeuAspProHisTrpValLeuThrAla 241
DB 670 TACGACAAACAGCAGCTGTGGAGGGAGCATCTCGACCCCACTGGCTCCTCAGCGCA 729
QY 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
DB 730 GCCCACTGCTTCAGAAACATACCGATGTGTCACTGGAGGTGGGGAGGCTCAGAC 789
QY 262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleIleGluPheAsnPro 281
DB 790 AAACCTGGGAGCTTCCCATCTCCCTGTGGCCAAAGATCATCATTTGAATTCACCC 849
QY 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
DB 850 ATGTACCCCAAGACATGACATCGCCCTCATGAAGCTGAGTCCACTCACTTTCTCA 909
QY 302 GlyThrValArgProIleCysLeuProPhePheAspGluLeuThrProAlaThrPro 321
DB 910 GGCACAGTCAGGCCATCTGTCTGCCCTCTTTGATGAGGAGCTCACTCCAGGCCACCCCA 969
QY 322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu 341

DB 970 CTCCTGATCATTTGGTGGGCTTTACGAGCAGAAATGAGGGAAGATGCTTGACATAC 1029
QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTrpGln 361
DB 1030 CTGACAGGCGTCACTCCAGGTCAATGACAGCACACGGTGCATGACAGCATGCGTACCAG 1089
QY 362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys 381
DB 1090 GGGAGAGTCCACGAGAGATGATGTGCGAGGCATCCCGAAGGGGTGTGACACCTGC 1149
QY 382 GlnGlyAspSerGlyGlyProLeuMetTrpGlnSerAspGlnTrpHisValValGlyIle 401
DB 1150 CAGGTCGACAGTGTGGGCCCCCTGATGTACCATCTGACAGTGGCATGTGGTGGGATC 1209
QY 402 ValSerTrpGlyTrpGlyCysGlyGlyProSerThrProGlyValTrpThrLysValSer 421
DB 1210 GTTAGCTGGGCTATGCTGCGGGGCCCGAGCACCAGAGTATACCAAGGTCTCA 1269
QY 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
DB 1270 GCCTATCTCACTGGATCTACAATGCTTGAAGGCTGAGCTG 1311

RESULT 10

US-10-295-027-133
; Sequence 133, Application US/10295027
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevesi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE OF INVENTION: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 133
; LENGTH: 1314
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-295-027-133

Alignment Scores:
Pred. No.: 0
Score: 2337.00
Length: 1314
Matches: 434

Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	99.79%	Indels:	0
DB:	44	Gaps:	0

US-10-803-530-2 (1-435) x US-10-295-027-133 (1-1314)

Qy	2	AspProIspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysProArg	21
Db	10	GATCCTGACAGTGATCAACTCTGAACACGCTCGATGTCAAAACCCCTCGCGAAACCCCGT	69
Qy	22	IleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSerLeu	41
Db	70	ATCCCCATGGAGACCTTCAGAAAGGTGGGATCCCCATCATCATAGCACTACTGAGCCGTG	129
Qy	42	AlaSerIleIleIleValValValLeuIleLysValIleLeuAspLysTyrTyrPheLeu	61
Db	130	CGGAGTATCATATGTGTGTCTCATCAAGGTGATTCGGATAAATACTACTTCTCTC	189
Qy	62	CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys	81
Db	190	TGCGGACGACCTCTCCACTTCACCGAGGAAGCAGCTGTGTGACGGAGAGCTGACTGT	249
Qy	82	ProLeuGlyGluAspGluHisCysValLysSerPheProGluGlyProAlaValAla	101
Db	250	CCCTTGGGGAGGACGAGAGACACTGTGTCAAGAGCTTCCCCGAAGGGGCTCGAGTGCGCA	309
Qy	102	ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp	121
Db	310	GTCGCGCTCTCAAGGACCGATCCACACTGACAGGTGTGGACTCGGCCACAGGAACTGG	369
Qy	122	PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet	141
Db	370	TTCCTGCGCTGTTCGACAACTTTCACAGAACTCTCGCTGAGACAGCCTGTAGGCAGATG	429
Qy	142	GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp	161
Db	430	GGCTACAGCAGCAAAACCACTTTCAGAGCTGTGGAGATTGGGCCACAGCAGATCTGGAT	489
Qy	162	ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys	181
Db	490	GTTGTTGAAATCACAGAAAAACAGCCAGGAGCTTCGATGCGGAATCAAGTGGGCGCTGT	549
Qy	182	LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro	201
Db	550	CTCTCAGGCTCCCTGGTCTCCCTGCGACTGTCTTGCTGTGGGAAGAGCCTGAAGACCCCC	609
Qy	202	ArgValValGlyGlyGluAlaSerValAspSerTrpProTrpGlnValSerIleGln	221
Db	610	CCTGTGGTGGGTGGGAGGAGGAGCCCTGTGTGGATTCTTGGCCCTTGGCAGGTCTCAG	669
Qy	222	TyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAla	241
Db	670	TACGCAAAACACACGCTGTGTGGAGGAGCATCTCTGGAACCCACATGGGTCTTCACGGCA	729
Qy	242	AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp	261
Db	730	GCCCACTGCTTCAGGAAACATACCGNATGTGTCAACTGGAAGGTGGCGCAGGCTCAGAC	789
Qy	262	LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleLeuGluPheAsnPro	281
Db	790	AAACTGGGCAGCTTCCCATCCCTGGGTGTGGCCAGATCATCATCTTGAATTCACACCCC	849
Qy	282	MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer	301
Db	850	ATGTATCCCCAAAGACAAATGACATCGGCCCTCATGAAGCTGCAGTTCACATCTTCTCA	909
Qy	302	GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro	321
Db	910	GGCAGAGTCAGGCCCATCTGTCTGCCCTTCTTTGATGAGGAGCTCACTCAGGCACCCCCA	969
Qy	322	LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu	341

Db	970	CTCTGGATCATTCGATGTGGGCTTTACGAGCAGCAATGGAGGGAGAGATGTCTGACATACTG	1029
Qy	342	LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGln	361
Db	1030	CTGCAGSCGTCAGTCCACGGTCATTGACAGCACACGGTGCAAATGCAGACCATGCGTACCAG	1089
Qy	362	GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys	381
Db	1090	GGGGAAGTCAACGAGAAGATCATGTGTGACGGCATCTCCGGNAGGGGGTGTGGACACCTGC	1149
Qy	382	GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle	401
Db	1150	CAGGGTCACAGTGTGTGGGCCCTCGATGTACCAATCTGCACAGTGGCATGTGGTGGGGCATC	1209
Qy	402	ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSer	421
Db	1210	GTTAGCTGGGGCTATGTGCTCGGGGGCCCGAGCACCCAGGAGTATACACCAAGGTCTCA	1269
Qy	422	AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu	435
Db	1370	GCCTATCTCAACTGGATCTCAATGTCTGGAAGGCTGAGCTG	1311

RESULT 11

RESOUR. 11
 US-10-295-027-778
 ; Sequence 778, Application US/10295027
 ; GENERAL INFORMATION:
 ; APPLICANT: Afar, Daniel
 ; APPLICANT: Aziz, Natasha
 ; APPLICANT: Ginsberg, Wendy M.
 ; APPLICANT: Gish, Kurt C.
 ; APPLICANT: Glynn, Richard
 ; APPLICANT: Hevezi, Peter A.
 ; APPLICANT: Mack, David H.
 ; APPLICANT: Murray, Richard
 ; APPLICANT: Watson, Susan R.
 ; APPLICANT: Eos Biotechnology, Inc.
 ; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
 ; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
 ; FILE REFERENCE: 018501-012500US
 ; CURRENT APPLICATION NUMBER: US/10/295,027
 ; CURRENT FILING DATE: 2002-11-13
 ; PRIOR APPLICATION NUMBER: US 09/663,733
 ; PRIOR FILING DATE: 2000-09-15
 ; PRIOR APPLICATION NUMBER: US 60/350,666
 ; PRIOR FILING DATE: 2001-11-13
 ; PRIOR APPLICATION NUMBER: US 60/335,394
 ; PRIOR FILING DATE: 2001-11-15
 ; PRIOR APPLICATION NUMBER: US 60/332,464
 ; PRIOR FILING DATE: 2001-11-21
 ; PRIOR APPLICATION NUMBER: US 60/334,393
 ; PRIOR FILING DATE: 2001-11-29
 ; PRIOR APPLICATION NUMBER: US 60/340,376
 ; PRIOR FILING DATE: 2001-12-14
 ; PRIOR APPLICATION NUMBER: US 60/347,211
 ; PRIOR FILING DATE: 2002-01-08
 ; PRIOR APPLICATION NUMBER: US 60/347,349
 ; PRIOR FILING DATE: 2002-01-10
 ; PRIOR APPLICATION NUMBER: US 60/355,250
 ; PRIOR FILING DATE: 2002-02-08
 ; PRIOR APPLICATION NUMBER: US 60/356,714
 ; PRIOR FILING DATE: 2002-02-13
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 1386
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 778
 ; LENGTH: 1314
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-295-027-778

Alignment Scores:		
Pred. No.:	0	Length:
Score:	2337.00	Matches:
		1314
		434

Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 99.79% Indels: 0
 DB: 44 Gaps: 0

US-10-803-530-2 (1-435) x US-10-295-027-778 (1-1314)

QY 2 AspProAspSerAspGlnProLeuAenSerLeuAenValLysProLeuAenArgLysProArg 21
 DB 10 GATCCTGACAGTATCAACTCTGAACAGCCTCGATGTCAAACCCCTCGGCAAAACCCCGT 69

QY 22 IleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSerLeu 41
 DB 70 ATCCCATGAGACCTTCAGAAAGGTGGGATCCCATCATCATAGCACTACTAGACCTG 129

QY 42 AlaSerIleIleIleValValLeuLeuLeuValLysValLysValLysValLysValLys 61
 DB 130 GCGAGTATCATATTGTGTGCTTCTCATCAAGAGTGATTCGTGAATAAATACTACTTCTC 189

QY 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGlnLeuAspCys 81
 DB 190 TCGGGGAGCCTTCACCTTCATCCGAGAGAGCAGCTGTGTGACGAGAGCTGGACTGT 249

QY 82 ProLeuGlyGluAspGluHisCysValLysSerPheProGluGlyProAlaValAla 101
 DB 250 CCCTTGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTCCCGAAGGCGCTGCAGTGGCA 309

QY 102 ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAenSerAlaThrGlyAsnTrp 121
 DB 310 GTCCGCCCTCTCCAAAGGACGATCCACACTGCAGAGTGTGGACTCGGCCACAGGAACTGG 369

QY 122 PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
 DB 370 TTCTCTGCTGTTTGACAACTTCAAGAGCTTCCGCTGAGAGAGCTGTAGGAGATG 429

QY 142 GlyTrpSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
 DB 430 GCGTACAGCAGCAAAACCCACTTTCAGAGCTGTGGAGATTGGCCAGACCAAGATCTGGAT 489

QY 162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
 DB 490 GTTGTGAAATACAGAAACAGCCAGGAGCTTCGCATGCGGAACTCAAGTGGGCCCTGT 549

QY 182 LeuSerGlySerLeuValSerLeuHisCysValLeuAlaCysGlyLysSerLeuLysThrPro 201
 DB 550 CTCTCAGCCTCCCTGCTCTCCCTGCACCTGTTCGCTGTGGAGAGAGCTGAAACCCCC 609

QY 202 ArgValValGlyGlyGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
 DB 610 CGTGTGTGTGGTGGGAGAGGCTCTGTGATTTCTTGGCCTTGGCAGGTCCAGATCCAG 669

QY 222 TyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAla 241
 DB 670 TACGACAAACAGCAGCTGTGGAGGGAGCATCTGGACCCCACTGGGTCTCCAGGCA 729

QY 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
 DB 730 GCGGCTGCTTCAGAAACATACCGATGTTCAACTGGAGGTGGGGAGGCTCAGAC 789

QY 262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsnPro 281
 DB 790 AAACCTGGGAGCTTCCCATCCCTGGCTGTGGCCAAAGATCATCATTTGAATTCACCCC 849

QY 282 MetTrpProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
 DB 850 ATGTACCCCAAGACATGACATCGCCCTCATGAAGCTGCAGTTCCTCCACTCACTTCTCA 909

QY 302 GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro 321
 DB 910 GGCACAGTCAGGCCCATCTGTCTGCCCTTCTTGTATGAGAGGCTCATCTCCAGCACCCCA 969

QY 322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu 341

DB 970 CTCTGGATCATTTGGATGGGGCTTTACGAAGCAGAAATGGAGGAGAGATGTTCTGACACTG 1029

QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTrpGln 361
 DB 1030 CTGCAGGCGTCAGTCCAGGTCAATGACAGCACAGGTCGATGCAATGCAGAGATCGGTACCG 1089

QY 362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys 381
 DB 1090 GGGGAAGTCAACCGAAGAGATGATGTGTGACAGGATCCCGAAGGGGTGTGGACACCTGC 1149

QY 382 GlnGlyAspSerGlyGlyProLeuMetTyrClnSerAspGlnTrpHisValValGlyIle 401
 DB 1150 CAGGTGACAGTGTGTGGGCCCCCTGATGTACCAATCTCCAGTGGCATGTGTGTGGCATC 1209

QY 402 ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSer 421
 DB 1210 GTTAGCTGGGCTATGCTGGGGGGCCGAGACCCCGAGGATATACCAAGGTCTTCA 1269

QY 422 AlaTrpLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
 DB 1270 GCCTATCTCACTGGATCTACAATGTCTGGAAGCTGAGCTG 1311

RESULT 12

US-10-295-027-790
 ; Sequence 790, Application US/10295027
 ; GENERAL INFORMATION:
 ; APPLICANT: Afar, Daniel
 ; APPLICANT: Aziz, Natasha
 ; APPLICANT: Ginsberg, Wendy M.
 ; APPLICANT: Gish, Kurt C.
 ; APPLICANT: Glynn, Richard
 ; APPLICANT: Hevesi, Peter A.
 ; APPLICANT: Mack, David H.
 ; APPLICANT: Murray, Richard
 ; APPLICANT: Watson, Susan R.
 ; APPLICANT: Eos Biotechnology, Inc.
 ; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
 ; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
 ; FILE REFERENCE: 018501-012500US
 ; CURRENT APPLICATION NUMBER: US/10/295,027
 ; CURRENT FILING DATE: 2002-11-13
 ; PRIOR APPLICATION NUMBER: US 09/663,733
 ; PRIOR FILING DATE: 2000-09-15
 ; PRIOR APPLICATION NUMBER: US 60/350,666
 ; PRIOR FILING DATE: 2001-11-13
 ; PRIOR APPLICATION NUMBER: US 60/335,394
 ; PRIOR FILING DATE: 2001-11-15
 ; PRIOR APPLICATION NUMBER: US 60/332,464
 ; PRIOR FILING DATE: 2001-11-21
 ; PRIOR APPLICATION NUMBER: US 60/334,393
 ; PRIOR FILING DATE: 2001-11-29
 ; PRIOR APPLICATION NUMBER: US 60/340,376
 ; PRIOR FILING DATE: 2001-12-14
 ; PRIOR APPLICATION NUMBER: US 60/347,211
 ; PRIOR FILING DATE: 2002-01-08
 ; PRIOR APPLICATION NUMBER: US 60/347,349
 ; PRIOR FILING DATE: 2002-01-10
 ; PRIOR APPLICATION NUMBER: US 60/355,250
 ; PRIOR FILING DATE: 2002-02-08
 ; PRIOR APPLICATION NUMBER: US 60/356,714
 ; PRIOR FILING DATE: 2002-02-13
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 1386
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 790
 ; LENGTH: 1314
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-295-027-790

Alignment Scores: 0 Length: 1314
 Pred. No.: 2337.00 Matches: 434
 Score:

Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.79% Indels: 0
DB: 44 Gaps: 0

US-10-803-530-2 (1-435) x US-10-295-027-790 (1-1314)

```
Qy 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysProArg 21
Db 10 GATCCTGACAGTATCAACCTCTGAACAGCCTCGATGTCAAAACCCCTCGCGAAACCCCGT 69
Qy 22 IleProMetGluThrPheArgLysValGlyLeuProIleIleIleAlaLeuSerLeu 41
Db 70 ATCCCCATGGAGAGCTTCAGAAAGGTGGGATCCCATCATAGCATAGCTACTGAGCCTG 129
Qy 42 AlaSerIleIleValValValValLysValLysValLysValLysValLysValLys 61
Db 130 CGAGATATCATATGTTGTTCTCATCAAGGTGATCTGGATTAATACTACTTCTC 189
Qy 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys 81
Db 190 TCGGGCAGCTCTCCACTTCATCCGAGGAGCAGCTGTGTGAGGAGAGCTGGACTGT 249
Qy 82 ProLeuGlyGluAspGluGluHisCysValLysSerPheProGluGlyProAlaValAla 101
Db 250 CCCTTGGGGAGGAGCAGGAGCAGCTGTGTCAAGAGCTTCCCGAAGGCGCTGCAGTGCA 309
Qy 102 ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp 121
Db 310 GTCCGCCCTCTCCAAAGGACCGATCCACACTGCAGGTGTGGATCGGCCACAGGGAACCTGG 369
Qy 122 PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
Db 370 TTCTCTGCTGTTTCGAACTTCACAAAGCTTCGCTGAGACAGCCTGTAGGCGAGTG 429
Qy 142 GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
Db 430 GGCTACAGCAGCAACCCACTTTCAGAGCTGTGGAGATGGCCACAGACAGGATCTGGAT 489
Qy 162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
Db 490 GTTGTGTAATCACAGAAACAGCCAGGAGCTTCGCATGCGGAACCTCAAGTGGGCCCTGT 549
Qy 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro 201
Db 550 CTCTCAGGCTCCCTGCTCTCCCTGACCTGTCTTGGCTGTGGAGAGAGCTCAAGTGGGCCCTGT 549
Qy 202 ArgValValGlyGlyGluGluAlaSerValAspSerTrpTrpGlnValSerIleGln 221
Db 610 CGTGTGGTGGTGGGAGGAGCCCTCTGTGGATCTTGGCTTGGCAGGTGAGCATCCAG 669
Qy 222 TyrAspLysGlnHisValCysGlySerIleLeuAspProHisTrpValLeuThrAla 241
Db 670 TACGACAAACAGCAGCTCTGTGGAGGAGCATCTTGGACCCCATGGTCTCTCAGGCA 729
Qy 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
Db 730 GCCCACTGCTTCAGGAACATACCATGTGTTCATCTGGAAGGTGGGGCAGGCTCAGAC 789
Qy 262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleIleGluPheAsnPro 281
Db 790 AAATGGGCAGCTTCCATCCCTGCTGTGGCAGAGATCATCATTAATTAACCC 849
Qy 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db 850 ATGTACCCCAAGACAATGACATCCGCTCATGAAGCTGCAGTTCCTCACTTCTCA 909
Qy 302 GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro 321
Db 910 GGCAAGTACAGGCCCATCTGTCTGCTGCTTCTTGTATGAGAGCTCACTCCAGCCACCCA 969
Qy 322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu 341
```

```
Db 970 CTCTGGATCATTTGGATGGGCTTTACGACGAGATGGAGGAGAGATGCTGTGACATCTG 1029
Qy 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGln 361
Db 1030 CTGCAGGCGTCACTCCAGTTCATTGACAGCAGCAGCGTGCATGACAGATGCGTACCAG 1089
Qy 362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys 381
Db 1090 GGGGAAGTCCACCGAAGATGATGTGTGCGAGCATCCCGAAGGGGTGTGGACACCTGC 1149
Qy 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle 401
Db 1150 CAGGCTGACAGTGTGGGCCCTCATGTACCAATCTGACAGTGGCATGTGTGGGCATC 1209
Qy 402 ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValThrLysValSer 421
Db 1210 GTTAGCTGGGCTATGCTGGGGGGCCGAGCAGCAGGAGTATACACCAAGGTCTCA 1269
Qy 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db 1270 GCCTATCTCACTGGATCTACAATGCTCTGGAAGGCTGAGCTG 1311
```

RESULT 13

```
US-10-295-027-830
; Sequence 830, Application US/10295027
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 830
; TYPE: DNA
; LENGTH: 1314
; ORGANISM: Homo sapiens
US-10-295-027-830
```

Alignment Scores: 0 Length: 1314
Pred. No.: 2337.00 Matches: 434
Score:

Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 99.79% Indels: 0
 DB: 44 Gaps: 0

US-10-803-530-2 (1-435) x US-10-295-027-830 (1-1314)

Qy 2 AspProAspSerAspGlnProLeuAenSerLeuAspValLysProLeuArgLysProArg 21
 Db 10 GATCTGACAGTATCAACCTCTGAACAGCCTCGATGTCAAAACCCCTCGCAACCCCGT 69
 Qy 22 IleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSerLeu 41
 Db 70 ATCCCATGGAGACCTTCAGAAAGTGGGATCCCATCATCATAGCACTACTGAGCCTG 129
 Qy 42 AlaSerIleIleIleValValValLeuIleLysValIleLeuAspLysTyThrPheLeu 61
 Db 130 GCGAGTATCATATTGTGTCTCATCAAGGTGATCTGGATAAATACTACTTCTCTC 189
 Qy 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys 81
 Db 190 TCGGGGAGCCTCTCCATTCATCCGAGGAGAGCTGTGTGACGGAGAGCTGGACTGT 249
 Qy 82 ProLeuGlyGluAspGluHisCysValLysSerPheProGluGlyProAlaValAla 101
 Db 250 CCCCTGGGGAGGAGAGACACTGTGTCAAGAGCTTCCCGAAGGGCCCTGCAGTGGCA 309
 Qy 102 ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp 121
 Db 310 GTCCGCCCTCTCCAAAGGACCGATCCACACTGCAGGTGTGGACTCGGCCACAGGAACTGG 369
 Qy 122 PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
 Db 370 TTCTCTGCTGTTCGACAACCTTCACAGAAGCTCTCGCTGAGACAGCGCTGTAGCAGATG 429
 Qy 142 GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
 Db 430 GGCTACAGCAGCAACCCACTTTCAGAGCTGTGAGATTGGCCACAGACGATCTGGAT 489
 Qy 162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
 Db 490 GTTGTGAATATCACAGAAACAGCCAGGAGCTTCGCATGCGGAACCTCAAGTGGGCCCTGT 549
 Qy 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro 201
 Db 550 CTCTCAGGCTCCCTGGTCTCCCTGCACCTCTTCGCTGTGGGAAGAGCCTGAAGACCCCC 609
 Qy 202 ArgValValGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
 Db 610 CGTGTGTGGTGGGGAGGAGCCCTCTGTGGATTCTTGGCCCTTGGCAGGTGAGCATCCAG 669
 Qy 222 TyrAspLysGlnHisValCysGlySerIleLeuAspProHisIleTrpValLeuThrAla 241
 Db 670 TACGACAAACAGCAGCTGTGTGGAGGAGCATCTCGGACCCCACTGGGTCTCTCACGGCA 729
 Qy 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
 Db 730 GCCCAGCTCTTCAGAAACATACCCGATGTGTCAACTGGAAGGTGCGGGCAGGCTCAGAC 789
 Qy 262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsnPro 281
 Db 790 AAATCTGGCAGCTTCCCATCCCTGGCTGTGGCCCAAGATCATCATTTGAATTCACCC 849
 Qy 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
 Db 850 ATGTACCCAAAGACAATGACATCGCCCTCATGAAGCTGCAAGTTCCTCCACTTCTCTCA 909
 Qy 302 GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro 321
 Db 910 GGCACAGTCAGGCCCATCTGTCTCCCTTCTTATGAGGAGCTCACTCCAGCCACCCCA 969
 Qy 322 LeuTrpIleIleGlyTyrGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu 341

Db 970 CTCTGGATCATTTGGATGGGCTTTACGACGACAATGGAGGAGAGATGTCTGCATACTG 1029
 Qy 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrln 361
 Db 1030 CTGCAGGCGTCAGTCCAGGTCAITTCACAGCACCGGTCAATGCAGACGATGCTGCCAG 1089
 Qy 362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys 381
 Db 1090 GGGGAAGTCCACCGAAGAAGATGTGTGCGAGGATATCCCGAAGGGGTGTGGACACCTGC 1149
 Qy 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle 401
 Db 1150 CAGGTGACAGTGTGTGGGCCCTGTGTGTACCAATCTGCAGGCGCATGTGGTGGGATC 1209
 Qy 402 ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSer 421
 Db 1210 GTTAGCTGGGCTATGCTGCGGGGCGCCGAGCACCCGAGGAGTATACCAAGGTCTCA 1269
 Qy 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
 Db 1270 GCCTATCTCACTGAGTCTCAATGTCTGGAAGGCTGAGCTG 1311

RESULT 14

US-10-295-027-979
 ; Sequence 979, Application US/10295027
 ; GENERAL INFORMATION:
 ; APPLICANT: Afar, Daniel
 ; APPLICANT: Aziz, Natasha
 ; APPLICANT: Ginsberg, Wendy M.
 ; APPLICANT: Gish, Kurt C.
 ; APPLICANT: Glynn, Richard
 ; APPLICANT: Hevazi, Peter A.
 ; APPLICANT: Mack, David H.
 ; APPLICANT: Murray, Richard
 ; APPLICANT: Watson, Susan R.
 ; APPLICANT: Eos Biotechnology, Inc.
 ; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
 ; FILE OF INVENTION: Methods of Screening for Modulators of Cancer
 ; FILE REFERENCE: 018501-012500US
 ; CURRENT APPLICATION NUMBER: US/10/295,027
 ; CURRENT FILING DATE: 2002-11-13
 ; PRIOR APPLICATION NUMBER: US 09/663,733
 ; PRIOR FILING DATE: 2000-09-15
 ; PRIOR APPLICATION NUMBER: US 60/350,666
 ; PRIOR FILING DATE: 2001-11-13
 ; PRIOR APPLICATION NUMBER: US 60/335,394
 ; PRIOR FILING DATE: 2001-11-15
 ; PRIOR APPLICATION NUMBER: US 60/332,464
 ; PRIOR FILING DATE: 2001-11-21
 ; PRIOR APPLICATION NUMBER: US 60/334,393
 ; PRIOR FILING DATE: 2001-11-29
 ; PRIOR APPLICATION NUMBER: US 60/340,376
 ; PRIOR FILING DATE: 2001-12-14
 ; PRIOR APPLICATION NUMBER: US 60/347,211
 ; PRIOR FILING DATE: 2002-01-08
 ; PRIOR APPLICATION NUMBER: US 60/347,349
 ; PRIOR FILING DATE: 2002-01-10
 ; PRIOR APPLICATION NUMBER: US 60/355,250
 ; PRIOR FILING DATE: 2002-02-08
 ; PRIOR APPLICATION NUMBER: US 60/356,714
 ; PRIOR FILING DATE: 2002-02-13
 ; Remaining prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 1386
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 979
 ; LENGTH: 1314
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-295-027-979

Alignment Scores: 0 2337.00 Length: 1314
 Pred. No.: 0 Matches: 434
 Score:

```
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.79% Indels: 0
DB: 44 Gaps: 0

US-10-803-530-2 (1-435) x US-10-295-027-979 (1-1314)

Qy 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysProArg 21
Db 10 GATCCTGACAGTATCAACCTCTGACAGCCTCGATGTCAAAACCCCTCGCAACCCCGT 69
Qy 22 IleProMetGluThrPheArgLysValGlyLeuProIleIleIleAlaLeuSerLeu 41
Db 70 ATCCCATGGAGACCTTCAGAAAGGTGGGATCCCATCATCATAGCACTACTGAGCCTG 129
Qy 42 AlaSerIleIleValValValLeuIleLysValIleLeuAspLysTyrrPheLeu 61
Db 130 CGAGATATCATATTGTGTCTCATCAAGGTGATTCTGGATAATACTACTTCTCTC 189
Qy 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys 81
Db 190 TCGGGGAGCCTCTCCATTCATCCGAGGAGCAGCTGTGTGACGGAGAGCTGGACTGT 249
Qy 82 ProLeuGlyGluAspGlnGluHisCysValLysSerPheProGluGlyProAlaValAla 101
Db 250 CCCTTGGGGGAGGACGAGGACCTGTGTCAAGAGCTTCCCGAAGGGCCCTGCACTGGCA 309
Qy 102 ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp 121
Db 310 GTCCGCTCTCCAGGACCGATCCACACTGACAGGTGTGACTCGGCCACAGGAGACTGG 369
Qy 122 PheSerAlaCysPheAspAspPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
Db 370 TTCTCTGCTGTTCGACAACTTCACAGAGCTCTCGCTGACAGACGCTGTAGGCAGATG 429
Qy 142 GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
Db 430 GGCTACAGCAGCAACCCACTTTCAGAGCTGTGGAGATTGGCCAGACGAGATCTGGAT 489
Qy 162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
Db 490 GTTGTGTAATCAGCAAAACAGCCAGAGCTTCGCATCGGAACTCAAGTGGGCCCTGT 549
Qy 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro 201
Db 550 CTCTCAGGCTCCCTGCTCTCCCTGCACTGTCTTGTCTGTGGGAGAGCCTGGAAGACCCC 609
Qy 202 ArgValValGlyGlyGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
Db 610 COTGTGTGGTGGGAGAGGAGCCTCTGTGGATTCTTGGCCTTGGCAGGTTCAGCATCCAG 669
Qy 222 TyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAla 241
Db 670 TACGACAAACAGACAGTCTGTGGAGGAGCATCTCTGGACCCCATCTGGTCTCTCAGGCA 729
Qy 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
Db 730 GCCCACTGCTTCAGGAACATACCGATGTGTTCAACTGGAAGGTGGGGCAGGCTCAGAC 789
Qy 262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsnPro 281
Db 790 NAACTGGGAGCAGCTTCCATCCCTGCTGTGGCAGAGATCATCATCATTTGAATTCACCCC 849
Qy 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db 850 ATGTACCCCAAGACAATGACATCCGCTCATGAAGCTGCAGTCTCCACTCTACTTCTCA 909
Qy 302 GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro 321
Db 910 GGCACAGTCAGGCCCATCTGTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 969
Qy 322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu 341
```

```
Db 970 CTCTGGATCATTTGGTGGGGCTTTACGACGAGCAATGGAGGGAAGATGTCTGACATACTG 1029
Qy 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTrpGln 361
Db 1030 CTGAGGCGCTCAGTCCAGGTCAATGACAGCACACGGTGCATGACAGACGATGCGTACCAG 1089
Qy 362 GlyGluValThrGluLysMetMetCysAlaGlyLeuProGluGlyGlyValAspThrCys 381
Db 1090 GGGGAAGTCAACCGAGAAGATGTGTGCGAGCATCCCGAAGGGGTGTGGACACCTGC 1149
Qy 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyLe 401
Db 1150 CAGGTCACAGCTGTGTGGGCCCTCTGATGTACCAATCTGACCAAGTGCATGTGTGGGCATC 1209
Qy 402 ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSer 421
Db 1210 GTTAGCTGGGCTATGGCTGGGGGGCCGAGCACCAGGAGTATACACCAAGGTCTCA 1269
Qy 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db 1270 GCCTATCTCAACTGGATCTACAATGTCTGGAAGGCTGAGCTG 1311

RESULT 15
US-60-625-561-448
; Sequence 448, Application US/60625561
; GENERAL INFORMATION:
; APPLICANT: MCCAFFREY, Ian
; APPLICANT: DOMON, Bruno
; TITLE OF INVENTION: PANCREATIC CANCER TARGETS AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001557
; CURRENT APPLICATION NUMBER: US/60/625,561
; CURRENT FILING DATE: 2004-11-08
; NUMBER OF SEQ ID NOS: 586
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 448
; LENGTH: 2079
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-625-561-448

Alignment Scores:
Pred. No.: 0 Length: 2079
Score: 2337.00 Matches: 434
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.79% Indels: 0
DB: 82 Gaps: 0

US-10-803-530-2 (1-435) x US-60-625-561-448 (1-2079)

Qy 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysProArg 21
Db 217 GATCCTGACAGTATCAACCTCTGACAGCCTCGATGTCAAAACCCCTGGCAACCCCGT 276
Qy 22 IleProMetGluThrPheArgLysValGlyLeuProIleIleIleAlaLeuLeuSerLeu 41
Db 277 ATCCCATGGAGACCTTCAGAAAGGTGGGATCCCATCATCATAGCACTACTGAGCCTG 336
Qy 42 AlaSerIleIleValValValLeuIleLysValIleLeuAspLysTyrrPheLeu 61
Db 337 CGAGATATCATATTGTGTCTCATCAAGGTGATTTCTGGATAATACTACTTCTCTC 396
Qy 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys 81
Db 397 TCGGGGAGCCTCTCCATTCATCCGAGGAGAGCAGCTGTGTGACGGAGAGCTGACTGT 456
Qy 82 ProLeuGlyGluAspGluGluHisCysValLysSerPheProGluGlyProAlaValAla 101
Db 457 CCCTTGGGGGAGGACGAGGACCTGTGTCAAGAGCTTCCCGAAGGGCCTGCACTGGCA 516
Qy 102 ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp 121
```

Db 517 GTCCGCCCTCTCCAGGACCGATCCACACTGCAGGTGCTGGACTCGGCCACAGGGAAGTGG 576
QY 122 PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
Db 577 TTCTCTGCTGCTGTTTCGAACTTCACAGAACTCTCGCTGAGACAGCTGTAGGCAGATG 636
QY 142 GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
Db 637 GGCTACAGCAGCAAAACCCACTTTCAGAGCTGTGGAGATTGGCCACAGACAGGATCTGGAT 696
QY 162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
Db 697 GTTGTGAAATACAGAAAAACAGCAGAGCTTCGCATGCGAACTCAAGTGGGCCCTGT 756
QY 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro 201
Db 757 CTCTCAGGCTCCCTGGTCTCCCTGCCTCTTCCCTGTGGAGAGAGCTGAAGACCCCC 816
QY 202 ArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
Db 817 CGTGTGCTGGTGGGAGGAGGCTCTGTGGATTCTTGGCCTTGGCAGGTACAGCATCCAG 876
QY 222 TyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAla 241
Db 877 TAGCACAAACAGCACGCTGTGGAGGAGCATCTCGACCCCACTGGGTCTCTCAGGCA 936
QY 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
Db 937 GCCCACTGCTTCAGGAAACATACCGATGTGTCACTGGAGGTGGGGAGGCTCAGAC 996
QY 262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsnPro 281
Db 997 AAACCTGGCAGCTTCCCTGCTGCTGGCTGTGGCAAGATCATCATTTGAATTCACCC 1056
QY 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db 1057 ATGTACCCCAAGACATGACATCGCCCTCATGAAGCTGCAGTCCCACTCACTTTCTCA 1116
QY 302 GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro 321
Db 1117 GGCACAGTCAGGCCCATCTGCTGCTGCTCTTTGATGAGGAGCTCACTCCAGCCACCCCA 1176
QY 322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu 341
Db 1177 CTCTGGATCATTTGGATGGGCTTTACGAAGCAGAAATGAGGGAAGATGCTGACATACTG 1236
QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGln 361
Db 1237 CTGCAGGCTCAGTCCAGTCAATTGACAGCACAGGTGCAATGCAGACGATGCGTACCAG 1296
QY 362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys 381
Db 1297 GGGGAAGTCACCGAAGATGATGTGTCAGGCAATCCGGNAGGGGTGTGGACACCTGC 1356
QY 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle 401
Db 1357 CAGGTGACAGTGGTGGGCTTATGATGACAAATCTGACAGTGGCATGTGTGGGCATC 1416
QY 402 ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSer 421
Db 1417 GTTAGCTGGGCTATGGCTGGGGGGCCGAGCACCCAGAGATATACCAAGGTCTCA 1476
QY 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db 1477 GCCTATCTCACTGGATCTACAATGTCTGGAAGGCTGAGCTG 1518

Search completed: December 7, 2005, 00:56:55
Job time : 6365 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2005, 13:45:59 ; Search time 576 Seconds
(without alignments)
1043.664 Million cell updates/sec

Title: US-10-803-530-2

Perfect score: 2342

Sequence: 1 MDPDSQPLNSLDVLRKLPK.....VYTKVSAYLNWYVWKAEL 435

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 7861189 seqs, 138195077 residues

Total number of hits satisfying chosen parameters: 7861189

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA Main:*

- 1: /cgn2_6/ptodata/1/paa/PCTUS COMB.pcp.*
- 2: /cgn2_6/ptodata/1/paa/US056 COMB.pcp.*
- 3: /cgn2_6/ptodata/1/paa/US073 COMB.pcp.*
- 4: /cgn2_6/ptodata/1/paa/US074 COMB.pcp.*
- 5: /cgn2_6/ptodata/1/paa/US075 COMB.pcp.*
- 6: /cgn2_6/ptodata/1/paa/US076 COMB.pcp.*
- 7: /cgn2_6/ptodata/1/paa/US077 COMB.pcp.*
- 8: /cgn2_6/ptodata/1/paa/US078 COMB.pcp.*
- 9: /cgn2_6/ptodata/1/paa/US079 COMB.pcp.*
- 10: /cgn2_6/ptodata/1/paa/US080 COMB.pcp.*
- 11: /cgn2_6/ptodata/1/paa/US081 COMB.pcp.*
- 12: /cgn2_6/ptodata/1/paa/US082 COMB.pcp.*
- 13: /cgn2_6/ptodata/1/paa/US083 COMB.pcp.*
- 14: /cgn2_6/ptodata/1/paa/US084 COMB.pcp.*
- 15: /cgn2_6/ptodata/1/paa/US085 COMB.pcp.*
- 16: /cgn2_6/ptodata/1/paa/US086 COMB.pcp.*
- 17: /cgn2_6/ptodata/1/paa/US087 COMB.pcp.*
- 18: /cgn2_6/ptodata/1/paa/US088 COMB.pcp.*
- 19: /cgn2_6/ptodata/1/paa/US089 COMB.pcp.*
- 20: /cgn2_6/ptodata/1/paa/US090 COMB.pcp.*
- 21: /cgn2_6/ptodata/1/paa/US091 COMB.pcp.*
- 22: /cgn2_6/ptodata/1/paa/US092 COMB.pcp.*
- 23: /cgn2_6/ptodata/1/paa/US093 COMB.pcp.*
- 24: /cgn2_6/ptodata/1/paa/US094 COMB.pcp.*
- 25: /cgn2_6/ptodata/1/paa/US095 COMB.pcp.*
- 26: /cgn2_6/ptodata/1/paa/US096 COMB.pcp.*
- 27: /cgn2_6/ptodata/1/paa/US097 COMB.pcp.*
- 28: /cgn2_6/ptodata/1/paa/US098 COMB.pcp.*
- 29: /cgn2_6/ptodata/1/paa/US099 COMB.pcp.*
- 30: /cgn2_6/ptodata/1/paa/US100 COMB.pcp.*
- 31: /cgn2_6/ptodata/1/paa/US101 COMB.pcp.*
- 32: /cgn2_6/ptodata/1/paa/US102 COMB.pcp.*
- 33: /cgn2_6/ptodata/1/paa/US103 COMB.pcp.*
- 34: /cgn2_6/ptodata/1/paa/US104 COMB.pcp.*
- 35: /cgn2_6/ptodata/1/paa/US105 COMB.pcp.*
- 36: /cgn2_6/ptodata/1/paa/US106 COMB.pcp.*
- 37: /cgn2_6/ptodata/1/paa/US107 COMB.pcp.*
- 38: /cgn2_6/ptodata/1/paa/US108 COMB.pcp.*
- 39: /cgn2_6/ptodata/1/paa/US109 COMB.pcp.*
- 40: /cgn2_6/ptodata/1/paa/US110 COMB.pcp.*
- 41: /cgn2_6/ptodata/1/paa/US111 COMB.pcp.*
- 42: /cgn2_6/ptodata/1/paa/US112 COMB.pcp.*
- 43: /cgn2_6/ptodata/1/paa/US114 COMB.pcp.*

- 44: /cgn2_6/ptodata/1/paa/US600 COMB.pcp.*
- 45: /cgn2_6/ptodata/1/paa/US601 COMB.pcp.*
- 46: /cgn2_6/ptodata/1/paa/US602 COMB.pcp.*
- 47: /cgn2_6/ptodata/1/paa/US603 COMB.pcp.*
- 48: /cgn2_6/ptodata/1/paa/US604 COMB.pcp.*
- 49: /cgn2_6/ptodata/1/paa/US605 COMB.pcp.*
- 50: /cgn2_6/ptodata/1/paa/US606 COMB.pcp.*
- 51: /cgn2_6/ptodata/1/paa/US607 COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2342	100.0	435	1	PCT-US01-18568-2	Sequence 2, Appli
2	2342	100.0	435	1	PCT-US01-18568A-2	Sequence 2, Appli
3	2342	100.0	435	1	PCT-US02-09671-1578	Sequence 1578, Ap
4	2342	100.0	435	1	PCT-US02-09671-1597	Sequence 1597, Ap
5	2342	100.0	435	30	US-10-030-688-2	Sequence 2, Appli
6	2342	100.0	435	34	US-10-473-127-1578	Sequence 1578, Ap
7	2342	100.0	435	34	US-10-473-127-1597	Sequence 1597, Ap
8	2342	100.0	435	38	US-10-803-530-2	Sequence 2, Appli
9	2342	100.0	461	1	PCT-US04-15258-3	Sequence 3, Appli
10	2342	100.0	461	1	PCT-US04-20741-7	Sequence 7, Appli
11	2338	99.8	435	1	PCT-US02-09671-1596	Sequence 1596, Ap
12	2338	99.8	435	26	US-09-659-151-6	Sequence 6, Appli
13	2338	99.8	435	31	US-10-180-719-6	Sequence 6, Appli
14	2338	99.8	435	34	US-10-473-127-1596	Sequence 1596, Ap
15	2338	99.8	435	40	US-11-045-577-6	Sequence 6, Appli
16	2338	99.8	435	41	US-11-183-914-6	Sequence 6, Appli
17	2337	99.8	437	1	PCT-US02-09671-1581	Sequence 1581, Ap
18	2337	99.8	437	1	PCT-US02-09671-1586	Sequence 1586, Ap
19	2337	99.8	437	1	PCT-US02-09671-1601	Sequence 1601, Ap
20	2337	99.8	437	1	PCT-US02-09671-1602	Sequence 1602, Ap
21	2337	99.8	437	1	PCT-US02-19297-89	Sequence 89, Appli
22	2337	99.8	437	1	PCT-US04-21227-7	Sequence 7, Appli
23	2337	99.8	437	1	PCT-US04-38689-7	Sequence 7, Appli
24	2337	99.8	437	31	US-10-173-999-89	Sequence 89, Appli
25	2337	99.8	437	32	US-10-295-027-779	Sequence 779, Appli
26	2337	99.8	437	32	US-10-295-027-791	Sequence 791, Appli
27	2337	99.8	437	32	US-10-295-027-831	Sequence 831, Appli
28	2337	99.8	437	32	US-10-295-027-1196	Sequence 1196, Appli
29	2337	99.8	437	34	US-10-473-127-1581	Sequence 1581, Appli
30	2337	99.8	437	34	US-10-473-127-1586	Sequence 1586, Appli
31	2337	99.8	437	34	US-10-473-127-1601	Sequence 1601, Appli
32	2337	99.8	437	34	US-10-473-127-1602	Sequence 1602, Appli
33	2337	99.8	437	39	US-10-991-287-7	Sequence 7, Appli
34	2337	99.8	437	39	US-10-994-117-7	Sequence 7, Appli
35	2337	99.8	437	50	US-60-625-561-207	Sequence 207, Appli
36	2333	99.6	435	39	US-10-918-711-490	Sequence 490, Appli
37	2333	99.6	435	39	US-10-918-754-2200	Sequence 2200, Appli
38	2333	99.6	435	48	US-60-495-114-2200	Sequence 2200, Appli
39	2333	99.6	435	48	US-60-495-135-490	Sequence 490, Appli
40	2333	99.6	435	50	US-60-625-561-206	Sequence 206, Appli
41	2329	99.4	492	1	PCT-US02-09671-1579	Sequence 1579, Ap
42	2329	99.4	492	1	PCT-US02-09671-1598	Sequence 1598, Ap
43	2329	99.4	492	30	US-10-030-688-4	Sequence 4, Appli
44	2329	99.4	492	34	US-10-473-127-1579	Sequence 1579, Ap
45	2329	99.4	492	34	US-10-473-127-1598	Sequence 1598, Ap

ALIGNMENTS

RESULT 1
PCT-US01-18568-2
; Sequence 2, Application PC/TUS0118568
; GENERAL INFORMATION:
; APPLICANT: Darrow, Andrew L

Db 61 LCGQPLHFIIPRKQLCDGELDCPLGDEEHCVKSPFEGPAVAVRLSKDRSTTLQVLDSATGN 120
Qy 121 WFSACFDNFTALAEACQMGYSKPTFRAVEIGPDQDLDVVEITENSQELRMNNSGP 180
Db 121 WFSACFDNFTALAEACQMGYSKPTFRAVEIGPDQDLDVVEITENSQELRMNNSGP 180
Qy 181 CLSGSLVSLHCLACGSKLTPRVVGGEASVDSWPQVSIQYDKQHVCGGSLDHPHWLT 240
Db 181 CLSGSLVSLHCLACGSKLTPRVVGGEASVDSWPQVSIQYDKQHVCGGSLDHPHWLT 240
Qy 241 AAHCFRKHDTVFNWKRAGSKLGFPSLAVAKIIIEFNPMYPKNDIALMKLOPPLTF 300
Db 241 AAHCFRKHDTVFNWKRAGSKLGFPSLAVAKIIIEFNPMYPKNDIALMKLOPPLTF 300
Qy 301 SGTVRPICLPFFDEELTPATPLWIIIGWFTKQNGKMSDILLOASVQVIDSTRCNADDAY 360
Db 301 SGTVRPICLPFFDEELTPATPLWIIIGWFTKQNGKMSDILLOASVQVIDSTRCNADDAY 360
Qy 361 QGEVTEKMKACGIPGGVDTCQDGGGGLMYQSDQHVGVISWGYGCGGSPTPGVYTKV 420
Db 361 QGEVTEKMKACGIPGGVDTCQDGGGGLMYQSDQHVGVISWGYGCGGSPTPGVYTKV 420
Qy 421 SAYLNWIYVNWKAEL 435
Db 421 SAYLNWIYVNWKAEL 435

RESULT 4

PCT-US02-09671-1597

; Sequence 1597, Application PC/TUS0209671

; GENERAL INFORMATION:

; APPLICANT: Zycos Inc.

; TITLE OF INVENTION: TRANSLATIONAL PROFILING

; FILE REFERENCE: 08191-026W01

; CURRENT APPLICATION NUMBER: PCT/US02/09671

; CURRENT FILING DATE: 2002-03-28

; PRIOR APPLICATION NUMBER: 60/279,495

; PRIOR FILING DATE: 2001-03-28

; PRIOR APPLICATION NUMBER: 60/292,544

; PRIOR FILING DATE: 2001-05-21

; PRIOR APPLICATION NUMBER: 60/310,801

; PRIOR FILING DATE: 2001-08-08

; PRIOR APPLICATION NUMBER: 60/326,370

; PRIOR FILING DATE: 2001-10-01

; PRIOR APPLICATION NUMBER: 60/336,780

; PRIOR FILING DATE: 2001-12-04

; PRIOR APPLICATION NUMBER: 60/358,985

; PRIOR FILING DATE: 2002-02-20

; NUMBER OF SEQ ID NOS: 2041

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1597

; LENGTH: 435

; TYPE: PRT

; ORGANISM: Homo sapiens

PCT-US02-09671-1597

Query Match 100.0%; Score 2342; DB 1; Length 435;
Best Local Similarity 100.0%; Pred. No. 3.2e-220;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDPDSQPLNSLDVPLKRPIMPETFRKVGIPPIIIALLSLASIIIVVVLKIVLDKYYF 60
Db 1 MDPDSQPLNSLDVPLKRPIMPETFRKVGIPPIIIALLSLASIIIVVVLKIVLDKYYF 60
Qy 61 LCGQPLHFIIPRKQLCDGELDCPLGDEEHCVKSPFEGPAVAVRLSKDRSTTLQVLDSATGN 120
Db 61 LCGQPLHFIIPRKQLCDGELDCPLGDEEHCVKSPFEGPAVAVRLSKDRSTTLQVLDSATGN 120
Qy 121 WFSACFDNFTALAEACQMGYSKPTFRAVEIGPDQDLDVVEITENSQELRMNNSGP 180
Db 121 WFSACFDNFTALAEACQMGYSKPTFRAVEIGPDQDLDVVEITENSQELRMNNSGP 180
Qy 181 CLSGSLVSLHCLACGSKLTPRVVGGEASVDSWPQVSIQYDKQHVCGGSLDHPHWLT 240

Db 181 CLSGSLVSLHCLACGSKLTPRVVGGEASVDSWPQVSIQYDKQHVCGGSLDHPHWLT 240
Qy 241 AAHCFRKHDTVFNWKRAGSKLGFPSLAVAKIIIEFNPMYPKNDIALMKLOPPLTF 300
Db 241 AAHCFRKHDTVFNWKRAGSKLGFPSLAVAKIIIEFNPMYPKNDIALMKLOPPLTF 300
Qy 301 SGTVRPICLPFFDEELTPATPLWIIIGWFTKQNGKMSDILLOASVQVIDSTRCNADDAY 360
Db 301 SGTVRPICLPFFDEELTPATPLWIIIGWFTKQNGKMSDILLOASVQVIDSTRCNADDAY 360
Qy 361 QGEVTEKMKACGIPGGVDTCQDGGGGLMYQSDQHVGVISWGYGCGGSPTPGVYTKV 420
Db 361 QGEVTEKMKACGIPGGVDTCQDGGGGLMYQSDQHVGVISWGYGCGGSPTPGVYTKV 420
Qy 421 SAYLNWIYVNWKAEL 435
Db 421 SAYLNWIYVNWKAEL 435

RESULT 5

US-10-030-688-2

; Sequence 2, Application US/10030688

; GENERAL INFORMATION:

; APPLICANT: Merck Patent GmbH

; TITLE OF INVENTION: Seripancrin

; FILE REFERENCE: SeripancrinUHS

; CURRENT APPLICATION NUMBER: US/10/030,688

; CURRENT FILING DATE: 2002-01-14

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2

; LENGTH: 435

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-030-688-2

Query Match 100.0%; Score 2342; DB 30; Length 435;
Best Local Similarity 100.0%; Pred. No. 3.2e-220;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDPDSQPLNSLDVPLKRPIMPETFRKVGIPPIIIALLSLASIIIVVVLKIVLDKYYF 60
Db 1 MDPDSQPLNSLDVPLKRPIMPETFRKVGIPPIIIALLSLASIIIVVVLKIVLDKYYF 60
Qy 61 LCGQPLHFIIPRKQLCDGELDCPLGDEEHCVKSPFEGPAVAVRLSKDRSTTLQVLDSATGN 120
Db 61 LCGQPLHFIIPRKQLCDGELDCPLGDEEHCVKSPFEGPAVAVRLSKDRSTTLQVLDSATGN 120
Qy 121 WFSACFDNFTALAEACQMGYSKPTFRAVEIGPDQDLDVVEITENSQELRMNNSGP 180
Db 121 WFSACFDNFTALAEACQMGYSKPTFRAVEIGPDQDLDVVEITENSQELRMNNSGP 180
Qy 181 CLSGSLVSLHCLACGSKLTPRVVGGEASVDSWPQVSIQYDKQHVCGGSLDHPHWLT 240
Db 181 CLSGSLVSLHCLACGSKLTPRVVGGEASVDSWPQVSIQYDKQHVCGGSLDHPHWLT 240
Qy 241 AAHCFRKHDTVFNWKRAGSKLGFPSLAVAKIIIEFNPMYPKNDIALMKLOPPLTF 300
Db 241 AAHCFRKHDTVFNWKRAGSKLGFPSLAVAKIIIEFNPMYPKNDIALMKLOPPLTF 300
Qy 301 SGTVRPICLPFFDEELTPATPLWIIIGWFTKQNGKMSDILLOASVQVIDSTRCNADDAY 360
Db 301 SGTVRPICLPFFDEELTPATPLWIIIGWFTKQNGKMSDILLOASVQVIDSTRCNADDAY 360
Qy 361 QGEVTEKMKACGIPGGVDTCQDGGGGLMYQSDQHVGVISWGYGCGGSPTPGVYTKV 420
Db 361 QGEVTEKMKACGIPGGVDTCQDGGGGLMYQSDQHVGVISWGYGCGGSPTPGVYTKV 420
Qy 421 SAYLNWIYVNWKAEL 435
Db 421 SAYLNWIYVNWKAEL 435

RESULT 6

```
US-10-473-127-1578
; Sequence 1578, Application US/10473127
; GENERAL INFORMATION:
; APPLICANT: Zycos Inc.
; TITLE OF INVENTION: TRANSLATIONAL PROFILING
; FILE REFERENCE: 08191-026W01
; CURRENT APPLICATION NUMBER: US/10/473,127
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 60/279,495
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: 60/292,544
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 60/310,801
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/326,370
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: 60/336,780
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/358,985
; PRIOR FILING DATE: 2002-02-20
; NUMBER OF SEQ ID NOS: 2041
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1578
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-473-127-1578
```

```
Query Match      100.0%; Score 2342; DB 34; Length 435;
Best Local Similarity 100.0%; Pred. No. 3.2e-220;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDPDSQPLNSLDVKPLRPMPMETFRKVGPIIIALLSLASIIIVVVLKVIDKYFF 60
Db 1 MDPDSQPLNSLDVKPLRPMPMETFRKVGPIIIALLSLASIIIVVVLKVIDKYFF 60
Qy 61 LCGQPLHFIPIKQLCDGELDCPLGEDEEHCVKSPFEGPAVAVRLSKDRSTLQVLD SATGN 120
Db 61 LCGQPLHFIPIKQLCDGELDCPLGEDEEHCVKSPFEGPAVAVRLSKDRSTLQVLD SATGN 120
Qy 121 WFSACFDNFTEALAEATACQMGYSKPTFRAVEIGPDQDLDVVEITENSQELMRNSSGP 180
Db 121 WFSACFDNFTEALAEATACQMGYSKPTFRAVEIGPDQDLDVVEITENSQELMRNSSGP 180
Qy 181 CLSGSLVSLHCLACGSKLTPRVVGEEASVDSWPQVSIQYDKQHVCGGSIIDPHWVLT 240
Db 181 CLSGSLVSLHCLACGSKLTPRVVGEEASVDSWPQVSIQYDKQHVCGGSIIDPHWVLT 240
Qy 241 AAHCFRKHDTVDNWKVRAGSKLGSFPSPSLAVAKIIIEFNPMYPKNDIALMKLQFPPLTF 300
Db 241 AAHCFRKHDTVDNWKVRAGSKLGSFPSPSLAVAKIIIEFNPMYPKNDIALMKLQFPPLTF 300
Qy 301 SGTVRPCLPFFDEELTPATPLWIIIGWFTKONGGKMSDILLOASVQVIDSTRCNAD DAY 360
Db 301 SGTVRPCLPFFDEELTPATPLWIIIGWFTKONGGKMSDILLOASVQVIDSTRCNAD DAY 360
Qy 361 QGEVTEKMKCAGIPEGGVDTCCQDGGPLMYQSDQHWVGVISWGYCGGSPSTFGVYTKV 420
Db 361 QGEVTEKMKCAGIPEGGVDTCCQDGGPLMYQSDQHWVGVISWGYCGGSPSTFGVYTKV 420
Qy 421 SAYLNWIYNWKAEL 435
Db 421 SAYLNWIYNWKAEL 435
```

RESULT 7

```
US-10-473-127-1597
; Sequence 1597, Application US/10473127
; GENERAL INFORMATION:
; APPLICANT: Zycos Inc.
; TITLE OF INVENTION: TRANSLATIONAL PROFILING
```

```
; FILE REFERENCE: 08191-026W01
; CURRENT APPLICATION NUMBER: US/10/473,127
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 60/279,495
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: 60/292,544
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 60/310,801
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/326,370
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: 60/336,780
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/358,985
; PRIOR FILING DATE: 2002-02-20
; NUMBER OF SEQ ID NOS: 2041
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1597
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-473-127-1597
```

```
Query Match      100.0%; Score 2342; DB 34; Length 435;
Best Local Similarity 100.0%; Pred. No. 3.2e-220;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDPDSQPLNSLDVKPLRPMPMETFRKVGPIIIALLSLASIIIVVVLKVIDKYFF 60
Db 1 MDPDSQPLNSLDVKPLRPMPMETFRKVGPIIIALLSLASIIIVVVLKVIDKYFF 60
Qy 61 LCGQPLHFIPIKQLCDGELDCPLGEDEEHCVKSPFEGPAVAVRLSKDRSTLQVLD SATGN 120
Db 61 LCGQPLHFIPIKQLCDGELDCPLGEDEEHCVKSPFEGPAVAVRLSKDRSTLQVLD SATGN 120
Qy 121 WFSACFDNFTEALAEATACQMGYSKPTFRAVEIGPDQDLDVVEITENSQELMRNSSGP 180
Db 121 WFSACFDNFTEALAEATACQMGYSKPTFRAVEIGPDQDLDVVEITENSQELMRNSSGP 180
Qy 181 CLSGSLVSLHCLACGSKLTPRVVGEEASVDSWPQVSIQYDKQHVCGGSIIDPHWVLT 240
Db 181 CLSGSLVSLHCLACGSKLTPRVVGEEASVDSWPQVSIQYDKQHVCGGSIIDPHWVLT 240
Qy 241 AAHCFRKHDTVDNWKVRAGSKLGSFPSPSLAVAKIIIEFNPMYPKNDIALMKLQFPPLTF 300
Db 241 AAHCFRKHDTVDNWKVRAGSKLGSFPSPSLAVAKIIIEFNPMYPKNDIALMKLQFPPLTF 300
Qy 301 SGTVRPCLPFFDEELTPATPLWIIIGWFTKONGGKMSDILLOASVQVIDSTRCNAD DAY 360
Db 301 SGTVRPCLPFFDEELTPATPLWIIIGWFTKONGGKMSDILLOASVQVIDSTRCNAD DAY 360
Qy 361 QGEVTEKMKCAGIPEGGVDTCCQDGGPLMYQSDQHWVGVISWGYCGGSPSTFGVYTKV 420
Db 361 QGEVTEKMKCAGIPEGGVDTCCQDGGPLMYQSDQHWVGVISWGYCGGSPSTFGVYTKV 420
Qy 421 SAYLNWIYNWKAEL 435
Db 421 SAYLNWIYNWKAEL 435
```

RESULT 8

```
US-10-803-530-2
; Sequence 2, Application US/10803530
; GENERAL INFORMATION:
; APPLICANT: Darrow, Andrew L
; APPLICANT: Qi, Jain-shen
; APPLICANT: Andrade-Gordon, Patricia
; TITLE OF INVENTION: DNA encoding human serine protease D-G
; FILE REFERENCE: ORT-1273
; CURRENT APPLICATION NUMBER: US/10/803,530
; CURRENT FILING DATE: 2004-03-17
; PRIOR APPLICATION NUMBER: US/09/607,745
; PRIOR FILING DATE: 2000-06-30
```

```
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Homo sapiens
us-10-803-530-2

Query Match      100.0%; Score 2342; DB 38; Length 435;
Best Local Similarity 100.0%; Pred. No. 3.2e-220;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDPDSQPLNSLDVVKPLRKPRIPMETFRKVGIPITIIALLSLASIIIVVLIKVLIDKYFF 60
Db 1 MDPDSQPLNSLDVVKPLRKPRIPMETFRKVGIPITIIALLSLASIIIVVLIKVLIDKYFF 60

Qy 61 LCGQPLHFI PRKQICDGLDCPLGEDEBHCVKSPPEGPAVAVRLSKDRSTIQVLD SATGN 120
Db 61 LCGQPLHFI PRKQICDGLDCPLGEDEBHCVKSPPEGPAVAVRLSKDRSTIQVLD SATGN 120

Qy 121 WFSACFNFTEALAEACRQMGYSKPTFRAVEIGPDQDLDVVEITENSQELRMNNSGP 180
Db 121 WFSACFNFTEALAEACRQMGYSKPTFRAVEIGPDQDLDVVEITENSQELRMNNSGP 180

Qy 181 CLSGSLVSLHCLACGSKLKTFRVVGGEASVDSNWPQVSIQYDKQHVCGGSILDPHWLT 240
Db 181 CLSGSLVSLHCLACGSKLKTFRVVGGEASVDSNWPQVSIQYDKQHVCGGSILDPHWLT 240

Qy 241 AAHCFRKHTDVFNWKVRAGSKLKTFRVVGGEASVDSNWPQVSIQYDKQHVCGGSILDPHWLT 300
Db 241 AAHCFRKHTDVFNWKVRAGSKLKTFRVVGGEASVDSNWPQVSIQYDKQHVCGGSILDPHWLT 300

Qy 301 SGTVRPICLPFFDEBELTPATPLWIIGWFTKQNGKMSDILLQASVQVIDSTRCNADAY 360
Db 301 SGTVRPICLPFFDEBELTPATPLWIIGWFTKQNGKMSDILLQASVQVIDSTRCNADAY 360

Qy 361 QGEVTEKMKACAGIPGEGVDTCQDSGGPLMYQSDQHVHVGVSVGWYCGGSPSTPGVYTKV 420
Db 361 QGEVTEKMKACAGIPGEGVDTCQDSGGPLMYQSDQHVHVGVSVGWYCGGSPSTPGVYTKV 420

Qy 421 SAYLNWIYNNWKAEL 435
Db 421 SAYLNWIYNNWKAEL 435

RESULT 9
PCT-US04-15258-3
; Sequence 3, Application PC/TUS0415258
; GENERAL INFORMATION:
; APPLICANT: diadexus, Inc.
; APPLICANT: Pilkington, Glenn
; APPLICANT: Keller, Gilbert-Andre
; APPLICANT: Li, Wenlu
; APPLICANT: Corral, Laura
; APPLICANT: Simon, Iris
; TITLE OF INVENTION: Ovr115 Antibody Compositions and Methods of Use
; FILE REFERENCE: DEX-0484
; CURRENT APPLICATION NUMBER: PCT/US04/15258
; PRIOR FILING DATE: 2004-05-21
; PRIOR APPLICATION NUMBER: US 60/559,730
; PRIOR FILING DATE: 2004-04-05
; PRIOR APPLICATION NUMBER: US 60/471,068
; PRIOR FILING DATE: 2003-05-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 3
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
PCT-US04-15258-3

Query Match      100.0%; Score 2342; DB 1; Length 461;
Best Local Similarity 100.0%; Pred. No. 3.4e-220;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDPDSQPLNSLDVVKPLRKPRIPMETFRKVGIPITIIALLSLASIIIVVLIKVLIDKYFF 60
Db 1 MDPDSQPLNSLDVVKPLRKPRIPMETFRKVGIPITIIALLSLASIIIVVLIKVLIDKYFF 60

Qy 61 LCGQPLHFI PRKQICDGLDCPLGEDEBHCVKSPPEGPAVAVRLSKDRSTIQVLD SATGN 120
Db 61 LCGQPLHFI PRKQICDGLDCPLGEDEBHCVKSPPEGPAVAVRLSKDRSTIQVLD SATGN 120

Qy 121 WFSACFNFTEALAEACRQMGYSKPTFRAVEIGPDQDLDVVEITENSQELRMNNSGP 180
Db 121 WFSACFNFTEALAEACRQMGYSKPTFRAVEIGPDQDLDVVEITENSQELRMNNSGP 180

Qy 181 CLSGSLVSLHCLACGSKLKTFRVVGGEASVDSNWPQVSIQYDKQHVCGGSILDPHWLT 240
Db 181 CLSGSLVSLHCLACGSKLKTFRVVGGEASVDSNWPQVSIQYDKQHVCGGSILDPHWLT 240

Qy 241 AAHCFRKHTDVFNWKVRAGSKLKTFRVVGGEASVDSNWPQVSIQYDKQHVCGGSILDPHWLT 300
Db 241 AAHCFRKHTDVFNWKVRAGSKLKTFRVVGGEASVDSNWPQVSIQYDKQHVCGGSILDPHWLT 300

Qy 301 SGTVRPICLPFFDEBELTPATPLWIIGWFTKQNGKMSDILLQASVQVIDSTRCNADAY 360
Db 301 SGTVRPICLPFFDEBELTPATPLWIIGWFTKQNGKMSDILLQASVQVIDSTRCNADAY 360

Qy 361 QGEVTEKMKACAGIPGEGVDTCQDSGGPLMYQSDQHVHVGVSVGWYCGGSPSTPGVYTKV 420
Db 361 QGEVTEKMKACAGIPGEGVDTCQDSGGPLMYQSDQHVHVGVSVGWYCGGSPSTPGVYTKV 420

Qy 421 SAYLNWIYNNWKAEL 435
Db 421 SAYLNWIYNNWKAEL 435

PCT-US04-20741-7
; Sequence 7, Application PC/TUS0420741
; GENERAL INFORMATION:
; APPLICANT: diadexus, Inc.
; APPLICANT: Papkoff, Jackie
; APPLICANT: Pilkington, Glenn
; APPLICANT: Keller, Gilbert-Andre
; APPLICANT: Li, Wenlu
; APPLICANT: Corral, Laura
; APPLICANT: Simon, Iris
; APPLICANT: Kmet, Muriel
; APPLICANT: Tang, Jianwen
; TITLE OF INVENTION: Pro104 Antibody Compositions and Methods of Use
; FILE REFERENCE: DEX-0491
; CURRENT APPLICATION NUMBER: PCT/US04/20741
; CURRENT FILING DATE: 2004-07-06
; PRIOR FILING DATE: 2003-11-17
; PRIOR APPLICATION NUMBER: US 60/523,271
; PRIOR FILING DATE: 2003-11-17
; PRIOR APPLICATION NUMBER: US 60/485,346
; PRIOR FILING DATE: 2003-06-27
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 7
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US04-20741-7

Query Match      100.0%; Score 2342; DB 1; Length 461;
Best Local Similarity 100.0%; Pred. No. 3.4e-220;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDPDSQPLNSLDVVKPLRKPRIPMETFRKVGIPITIIALLSLASIIIVVLIKVLIDKYFF 60
Db 1 MDPDSQPLNSLDVVKPLRKPRIPMETFRKVGIPITIIALLSLASIIIVVLIKVLIDKYFF 60
```

QY 61 LCGPLHFI PRKQLCDGELDCPLGEDBEHCVKSPFPGPAVAVRLSKDRSTLQVLD SATGN 120
 DB 61 LCGPLHFI PRKQLCDGELDCPLGEDBEHCVKSPFPGPAVAVRLSKDRSTLQVLD SATGN 120
 QY 121 WFSACFNFTEALAEACRQMGYSKPTFRRAVEIGPDQDLVVEITENSQELMRNSSGP 180
 DB 121 WFSACFNFTEALAEACRQMGYSKPTFRRAVEIGPDQDLVVEITENSQELMRNSSGP 180
 QY 181 CLSGLSVLSHCLACGKSLKTPRVVGGBEASVDSWPMQVSIQYDKQHVCGGSIILDPHWLT 240
 DB 181 CLSGLSVLSHCLACGKSLKTPRVVGGBEASVDSWPMQVSIQYDKQHVCGGSIILDPHWLT 240
 QY 241 AAHCFRKHDTVFNKVRAGSKLGSFPLSAVAKIIIEFNPMYPKNDIALMKLOFPLTF 300
 DB 241 AAHCFRKHDTVFNKVRAGSKLGSFPLSAVAKIIIEFNPMYPKNDIALMKLOFPLTF 300
 QY 301 SGTVRPCLPFPFDEELTPATPLWIIWGFTKQNGKMSDILLQASVOVIDSTRCNADAY 360
 DB 301 SGTVRPCLPFPFDEELTPATPLWIIWGFTKQNGKMSDILLQASVOVIDSTRCNADAY 360
 QY 361 QGEVTERKMCAGIPEGGVDTCQSDSGPLMYQSDQHVHVGVISWGYCGGSPSTPGVYTKV 420
 DB 361 QGEVTERKMCAGIPEGGVDTCQSDSGPLMYQSDQHVHVGVISWGYCGGSPSTPGVYTKV 420
 QY 421 SAYLNWIYNNWKAEL 435
 DB 421 SAYLNWIYNNWKAEL 435

RESULT 11
 PCT-US02-09671-1596
 ; Sequence 1596, Application PC/TUS0209671
 ; GENERAL INFORMATION:
 ; APPLICANT: Zycos Inc.
 ; TITLE OF INVENTION: TRANSLATIONAL PROFILING
 ; FILE REFERENCE: 08191-026W01
 ; CURRENT APPLICATION NUMBER: PCT/US02/09671
 ; CURRENT FILING DATE: 2002-03-28
 ; PRIOR APPLICATION NUMBER: 60/279,495
 ; PRIOR FILING DATE: 2001-03-28
 ; PRIOR APPLICATION NUMBER: 60/292,544
 ; PRIOR FILING DATE: 2001-05-21
 ; PRIOR APPLICATION NUMBER: 60/310,801
 ; PRIOR FILING DATE: 2001-08-08
 ; PRIOR APPLICATION NUMBER: 60/326,370
 ; PRIOR FILING DATE: 2001-10-01
 ; PRIOR APPLICATION NUMBER: 60/336,780
 ; PRIOR FILING DATE: 2001-12-04
 ; PRIOR APPLICATION NUMBER: 60/358,985
 ; PRIOR FILING DATE: 2002-02-20
 ; NUMBER OF SEQ ID NOS: 2041
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1596
 ; LENGTH: 435
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 PCT-US02-09671-1596

Query Match 99.8%; Score 2338; DB 1; Length 435;
 Best Local Similarity 99.8%; Pred. No. 7.8e-220;
 Matches 434; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MDPDSQPLNSLDVKPLRKPRIPMETFRKVGIPITIIALLSLASIIIVWLKILDKYVF 60
 DB 1 MDPDSQPLNSLDVKPLRKPRIPMETFRKVGIPITIIALLSLASIIIVWLKILDKYVF 60
 QY 61 LCGPLHFI PRKQLCDGELDCPLGEDBEHCVKSPFPGPAVAVRLSKDRSTLQVLD SATGN 120
 DB 61 LCGPLHFI PRKQLCDGELDCPLGEDBEHCVKSPFPGPAVAVRLSKDRSTLQVLD SATGN 120
 QY 121 WFSACFNFTEALAEACRQMGYSKPTFRRAVEIGPDQDLVVEITENSQELMRNSSGP 180

DB 121 WFSACFNFTEALAEACRQMGYSKPTFRRAVEIGPDQDLVVEITENSQELMRNSSGP 180
 QY 181 CLSGLSVLSHCLACGKSLKTPRVVGGBEASVDSWPMQVSIQYDKQHVCGGSIILDPHWLT 240
 DB 181 CLSGLSVLSHCLACGKSLKTPRVVGGBEASVDSWPMQVSIQYDKQHVCGGSIILDPHWLT 240
 QY 241 AAHCFRKHDTVFNKVRAGSKLGSFPLSAVAKIIIEFNPMYPKNDIALMKLOFPLTF 300
 DB 241 AAHCFRKHDTVFNKVRAGSKLGSFPLSAVAKIIIEFNPMYPKNDIALMKLOFPLTF 300
 QY 301 SGTVRPCLPFPFDEELTPATPLWIIWGFTKQNGKMSDILLQASVOVIDSTRCNADAY 360
 DB 301 SGTVRPCLPFPFDEELTPATPLWIIWGFTKQNGKMSDILLQASVOVIDSTRCNADAY 360
 QY 361 QGEVTERKMCAGIPEGGVDTCQSDSGPLMYQSDQHVHVGVISWGYCGGSPSTPGVYTKV 420
 DB 361 QGEVTERKMCAGIPEGGVDTCQSDSGPLMYQSDQHVHVGVISWGYCGGSPSTPGVYTKV 420
 QY 421 SAYLNWIYNNWKAEL 435
 DB 421 SAYLNWIYNNWKAEL 435

RESULT 12
 US-09-659-151-6
 ; Sequence 6, Application US/09659151
 ; GENERAL INFORMATION:
 ; APPLICANT: Bandman, Olga
 ; Hillman, Jennifer L.
 ; Yue, Henry
 ; Guegler, Karl J.
 ; Corley, Neil C.
 ; Tang, Tom Y.
 ; Shah, Purvi
 ; TITLE OF INVENTION: HUMAN PROTEASE MOLECULES
 ; NUMBER OF SEQUENCES: 24
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Dr.
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/659,151
 ; FILING DATE: 11-Sep-2000
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 09/008,271
 ; FILING DATE: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Mohan-Peterson, Sheela
 ; REGISTRATION NUMBER: 41,201
 ; REFERENCE/DOCKET NUMBER: PF-0458 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 650-855-0555
 ; TELEFAX: 650-845-4166
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 435 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: COLN0T13
 ; CLONE: 1337018
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 6 :
 US-09-659-151-6

Query Match 99.8%; Score 2338; DB 26; Length 435;
Best Local Similarity 99.8%; Pred. No. 7.8e-220;
Matches 434; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

SEQUENCE CHARACTERISTICS:
LENGTH: 435 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: COLNOT13
CLONE: 137018
SEQUENCE DESCRIPTION: SEQ ID NO: 6 :

US-10-180-719-6

Query Match 99.8%; Score 2338; DB 31; Length 435;
Best Local Similarity 99.8%; Pred. No. 7.8e-220;
Matches 434; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPDSQPLNSLDVPLKPRIPMETFRKVGIPPIIIALLSLASIIIVVVLKVIDLKYYF 60
DB 1 MDPDSQPLNSLDVPLKPRIPMETFRKVGIPPIIIALLSLASIIIVVVLKVIDLKYYF 60
QY 61 LCGQPLHFTPRKQLCDGELDCPLGEDEHCHVKSPFEGPAVAVRLSKDRSTLOVLSATGN 120
DB 61 LCGQPLHFTPRKQLCDGELDCPLGEDEHCHVKSPFEGPAVAVRLSKDRSTLOVLSATGN 120
QY 121 WFSACFDNTEALAEATACQMGYSKPTFRAVEIGPDQDLVDVEITENSQELMRNNSGP 180
DB 121 WFSACFDNTEALAEATACQMGYSKPTFRAVEIGPDQDLVDVEITENSQELMRNNSGP 180
QY 181 CLSGSLVSLHCLACGSLKTPRVVGEEASVDSWPQVSIQYDKQHVCGSILDPHWLT 240
DB 181 CLSGSLVSLHCLACGSLKTPRVVGEEASVDSWPQVSIQYDKQHVCGSILDPHWLT 240
QY 241 AAHCFRKHTDVFNWVKRAGSKLGFPSLAVAKIIIEFNPMYPKNDIALMKLOPPLTF 300
DB 241 AAHCFRKHTDVFNWVKRAGSKLGFPSLAVAKIIIEFNPMYPKNDIALMKLOPPLTF 300
QY 301 SGTVRPCLPFFDEELTPATPLWIIIGWFTKQNGKMSDILLOASVQVIDSTRCNADAY 360
DB 301 SGTVRPCLPFFDEELTPATPLWIIIGWFTKQNGKMSDILLOASVQVIDSTRCNADAY 360
QY 361 QGEVTEKMKCAGIPGEGVDTCQDGGPLMYOSDQHVHVGIVSWGVCGBPSTPGVYTKV 420
DB 361 QGEVTEKMKCAGIPGEGVDTCQDGGPLMYOSDQHVHVGIVSWGVCGBPSTPGVYTKV 420
QY 421 SAYLWNWYNVWKAEL 435
DB 421 SAYLWNWYNVWKAEL 435

RESULT 13

US-10-180-719-6
Sequence 6, Application US/10180719
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
Hillman, Jennifer L.
Yue, Henry
Guegler, Karl J.
Corley, Neil C.
Tang, Tom Y.
Shah, Purvi
TITLE OF INVENTION: HUMAN PROTEASE MOLECULES
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESS: Incyte Pharmaceuticals, Inc.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/180,719
FILING DATE: 25-Jun-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/008,271
FILING DATE: 16-Jan-1998
ATTORNEY/AGENT INFORMATION:
NAME: Mohan-Peterson, Sheela
REGISTRATION NUMBER: 41,201
REFERENCE/DOCKET NUMBER: PF-0458 US
TELECOMMUNICATION INFORMATION:

RESULT 14

US-10-473-127-1596
Sequence 1596, Application US/10473127
GENERAL INFORMATION:
APPLICANT: Zycos Inc.
TITLE OF INVENTION: TRANSLATIONAL PROFILING
FILE REFERENCE: 08191-026W01
CURRENT APPLICATION NUMBER: US/10/473,127
CURRENT FILING DATE: 2003-09-26
PRIOR APPLICATION NUMBER: 60/279,495
PRIOR FILING DATE: 2001-03-28
PRIOR APPLICATION NUMBER: 60/292,544
PRIOR FILING DATE: 2001-05-21
PRIOR APPLICATION NUMBER: 60/310,801
PRIOR FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: 60/326,370
PRIOR FILING DATE: 2001-10-01
PRIOR APPLICATION NUMBER: 60/336,780
PRIOR FILING DATE: 2001-12-04
PRIOR APPLICATION NUMBER: 60/358,985
PRIOR FILING DATE: 2002-02-20
NUMBER OF SEQ ID NOS: 2041

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1596
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-473-127-1596

Query Match          99.8%; Score 2338; DB 34; Length 435;
Best Local Similarity 99.8%; Pred. No. 7.8e-220;
Matches 434; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPDSQPLNSLDVVKLRPRIPMETFRKVGIPPIIIALLSLASIIIVVLIKVLDKYFF 60
DB 1 MDPDSQPLNSLDVVKLRPRIPMETFRKVGIPPIIIALLSLASIIIVVLIKVLDKYFF 60
QY 61 LCQQLPHFIPRKQLCDGELDCPLGEDEEHCVKSFPEGPAPAVRLSKDRSTLQVLD SATGN 120
DB 61 LCQQLPHFIPRKQLCDGELDCPLGEDEEHCVKSFPEGPAPAVRLSKDRSTLQVLD SATGN 120
QY 121 WFSACFDNFTEALAEATACRQMGYSKPTFRAVEIGPDQDLVDVVEITENSQELMRNSSGP 180
DB 121 WFSACFDNFTEALAEATACRQMGYSKPTFRAVEIGPDQDLVDVVEITENSQELMRNSSGP 180
QY 181 CLSGSLVSLHCLACGSKSLKTPRVVGGEASVDSWPQVSIQYDKQHVCGGSILDPHWLT 240
DB 181 CLSGSLVSLHCLACGSKSLKTPRVVGGEASVDSWPQVSIQYDKQHVCGGSILDPHWLT 240
QY 241 AAHCFRKHTDVFNVKVRAGSKLGSFSPSLAVAKIIIEFNPMYPKNDIALMKLOPPLTF 300
DB 241 AAHCFRKHTDVFNVKVRAGSKLGSFSPSLAVAKIIIEFNPMYPKNDIALMKLOPPLTF 300
QY 301 SGTVRPILCPFFDEELTPATPLWIIIGWFTKONGKMSDILLQASVQVIDSTRCNAD DAY 360
DB 301 SGTVRPILCPFFDEELTPATPLWIIIGWFTKONGKMSDILLQASVQVIDSTRCNAD DAY 360
QY 361 QGEVTEKMKCAGIPGGVDTCQDGGPLMYQSDQHVHVGIVSWGYGCGSPSTPGVYTKV 420
DB 361 QGEVTEKMKCAGIPGGVDTCQDGGPLMYQSDQHVHVGIVSWGYGCGSPSTPGVYTKV 420
QY 421 SAYLNWIYVNVKAE 435
DB 421 SAYLNWIYVNVKAE 435

RESULT 15
US-11-045-577-6
; Sequence 6, Application US/11045577
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; Hillman, Jennifer L.
; Yue, Henry
; Guegler, Karl J.
; Corley, Neil C.
; Tang, Tom Y.
; Shah, Purvi
; TITLE OF INVENTION: HUMAN PROTEASE MOLECULES
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/11/045,577
; FILING DATE: 27-Jan-2005
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/09/659,151
; FILING DATE: 11-Sep-2000
; APPLICATION NUMBER: 09/008,271
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Mohan-Peterson, Sheila
; REGISTRATION NUMBER: 41,201
; REFERENCE/DOCKET NUMBER: PF-0458 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 435 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: COLANNOT13
; CLONE: 1337018
; SEQUENCE DESCRIPTION: SEQ ID NO: 6 :
US-11-045-577-6

Query Match          99.8%; Score 2338; DB 40; Length 435;
Best Local Similarity 99.8%; Pred. No. 7.8e-220;
Matches 434; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPDSQPLNSLDVVKLRPRIPMETFRKVGIPPIIIALLSLASIIIVVLIKVLDKYFF 60
DB 1 MDPDSQPLNSLDVVKLRPRIPMETFRKVGIPPIIIALLSLASIIIVVLIKVLDKYFF 60
QY 61 LCQQLPHFIPRKQLCDGELDCPLGEDEEHCVKSFPEGPAPAVRLSKDRSTLQVLD SATGN 120
DB 61 LCQQLPHFIPRKQLCDGELDCPLGEDEEHCVKSFPEGPAPAVRLSKDRSTLQVLD SATGN 120
QY 121 WFSACFDNFTEALAEATACRQMGYSKPTFRAVEIGPDQDLVDVVEITENSQELMRNSSGP 180
DB 121 WFSACFDNFTEALAEATACRQMGYSKPTFRAVEIGPDQDLVDVVEITENSQELMRNSSGP 180
QY 181 CLSGSLVSLHCLACGSKSLKTPRVVGGEASVDSWPQVSIQYDKQHVCGGSILDPHWLT 240
DB 181 CLSGSLVSLHCLACGSKSLKTPRVVGGEASVDSWPQVSIQYDKQHVCGGSILDPHWLT 240
QY 241 AAHCFRKHTDVFNVKVRAGSKLGSFSPSLAVAKIIIEFNPMYPKNDIALMKLOPPLTF 300
DB 241 AAHCFRKHTDVFNVKVRAGSKLGSFSPSLAVAKIIIEFNPMYPKNDIALMKLOPPLTF 300
QY 301 SGTVRPILCPFFDEELTPATPLWIIIGWFTKONGKMSDILLQASVQVIDSTRCNAD DAY 360
DB 301 SGTVRPILCPFFDEELTPATPLWIIIGWFTKONGKMSDILLQASVQVIDSTRCNAD DAY 360
QY 361 QGEVTEKMKCAGIPGGVDTCQDGGPLMYQSDQHVHVGIVSWGYGCGSPSTPGVYTKV 420
DB 361 QGEVTEKMKCAGIPGGVDTCQDGGPLMYQSDQHVHVGIVSWGYGCGSPSTPGVYTKV 420
QY 421 SAYLNWIYVNVKAE 435
DB 421 SAYLNWIYVNVKAE 435

Search completed: December 5, 2005, 14:04:27
Job time : 580 secs
```


Qy	182	LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro	201
Db	806	CTCTCAGGCTCCCTGGTCTCCCTGCACCTGCTTCCTGTGGGAAGACCTGGAAGACCCCC	865
Qy	202	ArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleGln	221
Db	866	CTGTGGTGGGTGTGGAGAGGGCTCTGTGGATTCTTGGCTTGGCAGGTACAGATCCAG	925
Qy	222	TyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAla	241
Db	926	TAGCACAAACAGACACGCTCTGGAGGGAGCATCTGGACCCCACTGGGTCTTCACGGCA	985
Qy	242	AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp	261
Db	986	GCCCACTGCTTCAGGAACAATACCGATGTGTTCAACTGGAAGGTGGGGCAGGCTCAGAC	1045
Qy	262	LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleLeuGluPheAsnPro	281
Db	1046	AACTGGCAGCTTCCCATCTCCCTGGCTGTGGCCAGATCATCATCATTTGNAATCAACCCC	1105
Qy	282	MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer	301
Db	1106	ATGTACCCCAAAGACATGATCGCCCTCATGAAGCTGCAGTTCACCATCACTTCTCA	1165
Qy	302	GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro	321
Db	1166	GGCACAGTCAGGCCCATCTGCTGCTCCCTCTTTGATGAGGAGCTCATCTCAGACCAACCCCA	1225
Qy	322	LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu	341
Db	1226	CTCTGGATCATGTGATGGGGCTTTACGACGAGATGGAGGGAAGATGCTCTGACATCTG	1285
Qy	342	LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGln	361
Db	1286	CTCAGGGCGTCAGTCCAGGTCATTGACAGACACACGGTGCAATGCAGACGATGCGTACCAG	1345
Qy	362	GlyGluValThrGluLysMetMetCysAlaGlyLeProGluGlyValAspThrCys	381
Db	1346	GGGGAAGTCACCGAAGAGATGATGTGTGCAGGCATCCCGAAGGGGTGTGGACACCTGC	1405
Qy	382	GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle	401
Db	1406	CAGGGTGACAGTGTGGGCCCTGATGTACAATCTGACAGGTGGCATGTGGTGGGGATC	1465
Qy	402	ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSer	421
Db	1466	GTTAGTTGGGCTATGCTCGGGGGGCCGAGCACCCCGAGAGTATATACCAAGGTTCTCA	1525
Qy	422	AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu	435
Db	1526	GCCTATCTCAACTGGATCTCAATGTCTGGAAGGCTGAGCTG	1567

RESULT 3

```

US-10-206-921A-329
; Sequence 329, Application US/10206921A
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: 39780-34301C515
; CURRENT APPLICATION NUMBER: US/10/206,921A
; PRIORITY FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: US 19/052,586

```

	; PRIOR FILING DATE: 2002-01-15		
	; PRIOR APPLICATION NUMBER: PCT/US01/06520		
	; PRIOR FILING DATE: 2001-02-28		
	; PRIOR APPLICATION NUMBER: US 60/232,887		
	; PRIOR FILING DATE: 2000-09-15		
	; NUMBER OF SEQ ID NOS: 612		
	; SEQ ID NO 329:		
	LENGTH: 2063		
	TYPE: DNA		
	ORGANISM: Homo Sapien		
	US-10-206-921A-329		
	Alignment Scores:		
Pred. No.:	1..12e+233	Length:	2063
Score:	2297.50	Matches:	429
Percent Similarity:	98.85%	Conservative:	0
Best Local Similarity:	98.85%	Mismatches:	0
Query Match:	98.10%	Indels:	5
DB:	6	Gaps:	1
	US-10-803-530-2 (1-435) x US-10-206-921A-329 (1-2063)		
QY	2 AspProAspSerAspGlnProLeuAasnSerLeuAspValLysProLeuArgLysProArg	21	
Db	219 GATCCTGACAGTGATCAACTCTGAACAGCCCTCGATGTCAAACCCTCTGGCAAAACCCCCTGC	278	
QY	22 IleProMetGluThrPheArgLyseValcIlyleProfilellelleAlaLeuLeuSerLeu	41	
Db	279 ATCCCCATGAGACCTTCAGAAAGTGGGGATCCCCATCATCATAGCACCTACTGAGCCGTG	338	
QY	42 AlaSerllellelValvalValleullelysVallilleuleuApplvsTyrrPheleu	61	
Db	339 GCGAGTATCATCATTTGGTGTCTCATCAAGGTGATTCTGGATAAAATACTTACTTTCCTC	398	
QY	62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys	81	
Db	399 TCGGGGAGGCCCTCTCCACTTCATCCGAGAAGCAGCTGTGTGACGGAGAGCTGGACTGT	458	
QY	82 ProLeuGlyGluAspGluGluHisCysVallysserPheProGluGlyProAlaValAla	101	
Db	459 CCCTTTGGGGAGGACGAGGAGCAGCTGTGTCAAGAGCTTCCCGAAGGGCCTGCAGTGCCA	518	
QY	102 ValArgLeuSerLysAspArgserThrlieuGlnValLeuAspSerAlaThrCllyAsnTrp	121	
Db	519 GTCCGCCTCTCCAAAGGACCGATCCACACTTGCAGGTGTGGACTCGGCCACACGGSAAC	578	
QY	122 PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet	141	
Db	579 TTCTCTGCCTGTTTCGACAACCTTCACAGAACCTCTCGCTGAGACAGCCTGTAGGCAGATG	638	
QY	142 GlyTyrSerSertysProThrPheArgAlaValGluIlleGlyProAspGlnAspLeuAsp	161	
Db	639 GGCTACAGC-----AGAGCTGTGGAGATTGGCCACAGACAGGATCTGGAT	683	
QY	162 ValValGluIlleThrGluAasnerGlnGluLeuArgMetArgAasnerSerGlyProCys	181	
Db	684 GTTGTGAAAATCACAGAAAAACAGCAGGAGCTTCGCATCGCGGAATCTCAAGTCGGGCCCTGT	743	
QY	182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysserLeuLysThrPro	201	
Db	744 CTCTCAGGCTCCCTGGHTCTCCTGCACITGTCTTGCTGTGGGAAGAGCCTGAAGACCCCC	803	
QY	202 ArgValValGlyGlyGluGlualaserValAspSerTrpProTrpGlnValSerlleGln	221	
Db	804 CGTGTGGTGGGTGGGAGGAGGCCCTCTGTGGATTCTTGGCCTTTGGCAGTCAGCATCCAG	863	
QY	222 TyrAspLysGlnHissValCyseGlySerlleLeuAspProHisTrpvallleuthrAla	241	
Db	864 TACGACAAAACAGCAGCTGTGTGGAGGAGCATCTCTGGAACCCCATCTGGGTCTCTACGGCA	923	
QY	242 AlahisCysPheArgLyseHisThrAspValPheAsnTrpLysValArglaGlySerApp	261	
Db	924 GCCCACTGCTTCAGGAAACATACCGATGTGTCAACTGGAAGGTGCGGAGGCTCAGAC	983	


```

US-10-868-184C-2378
; Sequence 2378, Application US/10868184C
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS805
; CURRENT APPLICATION NUMBER: US/10/868,184C
; CURRENT FILING DATE: 2004-06-16
; PRIOR APPLICATION NUMBER: 60/278,650
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 09/833,245
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: PCT/US01/11988
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: PCT/US00/06043
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06012
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06058
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06044
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06059
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06042
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06014
; PRIOR FILING DATE: 2000-03-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 13046
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2378
; LENGTH: 1082
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-868-184C-2378

Alignment Scores:
Pred. No.: 4,378-42 Length: 1082
Score: 478.00 Matches: 101
Percent Similarity: 55.81% Conservative: 48
Best Local Similarity: 37.83% Mismatches: 86
Query Match: 20.41% Indels: 32
DB: 6 Gaps: 8

US-10-803-530-2 (1-435) x US-10-868-184C-2378 (1-1082)
QY 178 SerGlyProCysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSer 197
DB 97 TCAGGACCA-----TGCGGCGGACGG 117
QY 198 LeuLysThrProArgValValGlyGluGluAlaSerValAspSerTrpProTrpGln 217
DB 118 GTCAATCAGTCGCGCATCGTGGTGGAGAGAGCGCGAACTCGGGCGTGGCGGTGGCAG 177
QY 218 ValSerIleGlnTyraAspLysGlnHisValCysGlySerIleLeuAspProHisTrp 237
DB 178 GGGAGCTCGCCCTGGGATCCACGATGCGGAGTGAGCTGCGGCTCGACCCACCGCTGG 237
QY 238 ValLeuThrAlaAlaHisCysPheArgLysHisThrAspValPheAsn-----Tip 254
DB 238 GCACTCAGCGGGCGCACTGCTTTGAAACCTATAGTACCTTAGTGATCCCTCGGGTGG 297
QY 255 LysValArgAlaGlySerAspLysLeuGlySerPheProSerLeuAlaValAlaLys--- 273
DB 298 ATGCTCAGTTGGC-----CAGCTGACTTCCATGCCATCTCTGAGCGCTGCAGGCC 351
QY 274 -----IleIleIleIleGluPheAsnProMetTyraProLysAspAsn 287
DB 352 TACTACACCGCTTACTTCGATCGAATATTATCTGAGCCCTCGCTACCTGGGGAATTC 411
QY 288 -----AspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSerGlyThrValArg 305
DB 97 TCAGGACCA-----TGCGGCGGACGG 117

```

RESULT 7

US-10-206-921A-319

; Sequence 319, Application US/10206921A

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE OF INVENTION: ACIDS ENCODING THE SAME

; FILE REFERENCE: 39780-3430R1C515

; CURRENT APPLICATION NUMBER: US/10/206,921A

; CURRENT FILING DATE: 2002-07-26

; PRIOR APPLICATION NUMBER: US 10/052,586

; PRIOR FILING DATE: 2002-01-15

; PRIOR APPLICATION NUMBER: PCT/US01/06520

; PRIOR FILING DATE: 2001-02-28

; PRIOR APPLICATION NUMBER: US 60/232,887

; PRIOR FILING DATE: 2000-09-15

; NUMBER OF SEQ ID NOS: 612

; SEQ ID NO 319

; LENGTH: 2103

; TYPE: DNA

; ORGANISM: Homo Sapien

US-10-206-921A-319

Alignment Scores:

Pred. No.: 2,398-41 Length: 2103

Score: 475.00 Matches: 97

Percent Similarity: 56.33% Conservative: 41

Best Local Similarity: 39.59% Mismatches: 97

Query Match: 20.28% Indels: 10

DB: 6 Gaps: 5

US-10-803-530-2 (1-435) x US-10-206-921A-319 (1-2103)

```
QY 190 HisCysLeuAlaCysGly-----LysSerLeu---LysThrProArgValVal 204
Db 557 CATTTGC-----TGGCGAACAGAAAGTAAACTCTAGGTCAGAGTCTCAGGATCGTT 610
QY 205 GlyGlyGluGluAlaSerValAspSerTrpProTtpGlnValSerIleGlnTyrAspLys 224
Db 611 GGTGGACAGAGTAAAGAGGGTGAATGGCCCTGGCAGGCTAGCTGCAGTGGGATGGG 670
QY 225 GlnHisValCysGlySerIleLeuAspProHisTrpValLeuThrAlaAlaHisCys 244
Db 671 AGTCATCGCTGGGACCAACCTTAATAATGACCAATGCTGTGAGTGTCTGCTCACTGT 730
QY 245 PheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAspLysLeuGly 264
Db 731 TTTCAACATATAAGAACCTCCAGATGCTGCTCTCTTGGAGTAAACAATAAAACCT 790
QY 265 SerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsnProMetTyrPro 284
Db 791 TCGAAATGAACGGGGTCTCGGAGATAATTTGTCATGAAATAACAAACCCATCA 850
QY 285 LysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSerGlyThrVal 304
Db 851 CATGACTATGATATTCTCTTCAGAGCTTTCTAGCCCTGTTCCCTACACAAATGCAGTA 910
QY 305 ArgProIleCysLeuProPheAspGluGluLeuThrProAlaThrProLeuTrpIle 324
Db 911 CATAGAGTTTGTCTCCCTGATGCATCTATGAGTTTCAACCAAGGTGATGTGTTGTG 970
QY 325 IleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeuGlnAla 344
Db 971 ACAGAGTTTGGAGCACTGAAA---AATGATGTTACAGTCAAAATCATCTTCACACAGA 1027
QY 345 SerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGlnGlyGluVal 364
Db 1028 CAGGTGACTCTCATAGACCTACAACTTCAATGAACCTCAAGCTTACAATCAGCCATA 1087
QY 365 ThrGluLysMetCysAlaGlyIleProGluGlyGlyValAspThrCysGlnGlyAsp 384
Db 1088 ACTCCTAGAATGTTATGTCTGCTGGCTCCTTAGAAGGAAACACAGATGCATGCCAGGTGAC 1147
QY 385 SerGlyGlyProLeuMetTyrGlnSer-----AspGlnTrpHisValValGlyIleVal 402
Db 1148 TCTGGAGGACCACTGTTAGTTAGTTCAGATGCTAGAGATATCTGGTACCCTCTGGAATAGT 1207
QY 403 SerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSerAla 422
Db 1208 ACTGGGGAGATGAATGTGCGAAACCCACACAGCCCTGGTGTATTACTAGATTACGGCC 1267
QY 423 TyrLeuAsnTrpIle 427
Db 1268 TTGCGGGAGCTGGATT 1282

RESULT 8
US-11-265-762-145
; Sequence 145, Application US/11265762
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Sean A.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-227-999
; CURRENT APPLICATION NUMBER: US/11/265,762
; CURRENT FILING DATE: 2005-11-02
; PRIOR APPLICATION NUMBER: US/09/796,753
; PRIOR FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 09/183,175
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 09/223,094
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/223,546
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/224,246
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/259,388
; PRIOR FILING DATE: 1999-02-26
```

```
; PRIOR APPLICATION NUMBER: 60/122,458
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: 09/312,359
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 09/336,536
; PRIOR FILING DATE: 1999-06-18
; PRIOR APPLICATION NUMBER: 09/342,687
; PRIOR FILING DATE: 1999-06-29
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 162
; SEQ ID NO 145
; LENGTH: 5058
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-265-762-145

Alignment Scores:
Pred. No.: 8,588-41 Length: 5058
Score: 475.00 Matches: 97
Percent Similarity: 56.33% Conservative: 41
Best Local Similarity: 39.59% Mismatches: 97
Query Match: 20.28% Indels: 10
DB: 7 Gaps: 5

US-10-803-530-2 (1-435) x US-11-265-762-145 (1-5058)
QY 190 HisCysLeuAlaCysGly-----LysSerLeu---LysThrProArgValVal 204
Db 566 CATTTGC-----TGGCGAACAGAAAGTAAACTCTAGGTCAGAGTCTCAGGATCGTT 619
QY 205 GlyGlyGluGluAlaSerValAspSerTrpProTtpGlnValSerIleGlnTyrAspLys 224
Db 620 GGTGGACAGAGTAAAGAGGGTGAATGGCCCTGGCAGGCTAGCTGCAGTGGGATGGG 679
QY 225 GlnHisValCysGlySerIleLeuAspProHisTrpValLeuThrAlaAlaHisCys 244
Db 680 AGTCATCGCTGGGAGCACTTAATAATGCCATGCTGGTGTGAGTGTCTGCTCACTGT 739
QY 245 PheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAspLysLeuGly 264
Db 740 TTTCAACATATAAGAACCTCCAGATGCTGCTCTCTTGGAGTAAACAATAAAACCT 799
QY 265 SerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsnProMetTyrPro 284
Db 800 TCGAAATGAACGGGGTCTCGGAGATAATTTGTCATGAAATAACAAACCCATCA 859
QY 285 LysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSerGlyThrVal 304
Db 860 CATGACTATGATATTCTCTTCAGAGCTTTCTAGCCCTGTTCCCTACACAAATGCAGTA 919
QY 305 ArgProIleCysLeuProPheAspGluGluLeuThrProAlaThrProLeuTrpIle 324
Db 920 CATAGAGTTTGTCTCCCTGATGCATCTATGAGTTTCAACCAAGGTGATGTGTTGTG 979
QY 325 IleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeuGlnAla 344
Db 980 ACAGAGTTTGGAGCACTGAAA---AATGATGTTTACAGTCAAAATCATCTTCGACAGCA 1036
QY 345 SerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGlnGlyGluVal 364
Db 1037 CAGGTGACTCTCATAGACCTACAACCTTGAATGAACCTCAAGCTTACAATCAGCCATA 1096
QY 365 ThrGluLysMetCysAlaGlyIleProGluGlyGlyValAspThrCysGlnGlyAsp 384
Db 1097 ACTCCTAGAATGTTATGTCTGGCTCCTTAGAAGGAAACACAGATGCATGCCAGGTGAC 1156
QY 385 SerGlyGlyProLeuMetTyrGlnSer-----AspGlnTrpHisValValGlyIleVal 402
Db 1157 TCTGGAGGACCACTGTTAGTTAGTTCAGATGCTAGAGATATCTGTTACTGGAATAGT 1216
QY 403 SerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSerAla 422
Db 1217 AGCTGGGAGATGAATGTGCGAAACCCACACAGCCCTGGTGTATTACTAGATTACGGCC 1276
```

```
Qy 423 TyrLeuAsnTrpIle 427
Db 1277 TTGGGAGTGGATT 1291

RESULT 9
US-11-268-554-707
; Sequence 707, Application US/11268554
; GENERAL INFORMATION:
; APPLICANT: DOMON, Bruno et al.
; TITLE OF INVENTION: PANCREATIC CANCER TARGETS AND USES
; FILE REFERENCE: CL001557
; CURRENT APPLICATION NUMBER: US/11/268,554
; CURRENT FILING DATE: 2005-11-08
; NUMBER OF SEQ ID NOS: 1004
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 707
; LENGTH: 3305
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-268-554-707

Alignment Scores:
Pred. No.: 8 52e-40 Length: 3305
Score: 463.00 Matches: 123
Percent Similarity: 48.21% Conservative: 65
Best Local Similarity: 31.54% Mismatches: 151
Query Match: 19.77% Indels: 52
DB: 7 Gaps: 15

US-10-803-530-2 (1-435) x US-11-268-554-707 (1-3305)
Qy 58 TyrTyrPheLeuCysGly-----GlnProLeuHisPheIleProArgLysGln 73
Db 1667 CACCAGTTCCAGTGCAGAAACAAGTTCTGCAAGCCCTCTTCTGGTGC----- 1714

Qy 74 LeuCysAspGlyGluLeuAspCysProLeuGlyGluAspGluGluHisCysValLysSer 93
Db 1715 ----TGGCAGGTGTGAACGACTGCGGAGACACACAGCGACGACGAGGGTGC----- 1762

Qy 94 PheProGluGlyProAlaValAlaValArgLeuSerLysAspArgSerThrLeuGlnVal 113
Db 1763 -----AGTTGTCGCGCCACGACCTTCAGGTGTTCCAAATGGAGTGC-----CTC 1807

Qy 114 LeuAspSerAlaThrGlyAsnTrpPheSerAlaCysPheAspAsnPheThrGluAlaLeu 133
Db 1808 TCGAAAAGCCAGCAGTGCATATGGGAAGCAGCAGTGTGGGACGGTCCGACGAGGCC--- 1864

Qy 134 AlaGluThrAlaCysArgGlnMetGlyTyr-----SerSerLysProThrPheArgAla 151
Db 1865 -----TCTGCCCCAAGGTGAACGTCGTCACTTGTACCAACACACACTACCGTGC 1915

Qy 152 ValGluIleGlyProAspGlnAspLeuAspValValGluIleThrGluAsnSerGlnGlu 171
Db 1916 -----CTCAATGGGCTCTGTTGACCAAGGCAACCCCTGAG 1951

Qy 172 LeuArgMetArgAsnSerSerGlyProCysLeuSerGlySer-LeuValSerLeuHisCy 191
Db 1952 TGT-----GACGGAGAGGAGGACTGTAGCGACGGCTCAGATGAGAAGGACTGCGAC 2002

Qy 191 sLeuAlaCysGlyLysSerLeuLysThrProArgValValGlyGlyGluGluAlaSerVa 211
Db 2003 TGTGGGTGCGGTTCATTACG-AGCAGGCTCGTGTGTTGGGGGACCGATCGGATGA 2061

Qy 211 lAspSerTrpProTrpGlnValSerIleGln---TyrAspLysGlnHisValCysGlyG1 230
Db 2062 GGGCAGTGGCCCTGGCAGGTAGCTGCTGCTGGGCGCAGGGCCACATCTGCGGTGC 2121

Qy 230 yserIleLeuAspProHisTrpValLeuThrAlaAlaHisCysPhe----- 245
Db 2122 TTCCCTCATCTCCCACTGGCTGGTCTCTGCGGCACACTGCTATCATGATGACAGAGG 2181

Qy 246 -ArgLysHisThrAspValPheAsnTrpLysValAlaAlaGly-----SerAspLysLe 263
Db 2182 ATTCAAGTACTCAGACCCACAGCAGTGCAGCGCTTCTCTGGCTTGCACGACGAGGCCA 2241

Qy 263 uGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsnProMetTy 283
Db 2242 GCGCAGCGCCCTGGGTGCGAGGCGCGCTCAAGCGCATCATCTCCACCCCTTCTT 2301

Qy 283 rProLys-----AspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPh 300
Db 2302 CAATGACTTCACTTCGACTATGATCGCGTCTGGAGCTGGAGAAACCGGACAGATA 2361

Qy 300 eSerGlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaTh 320
Db 2362 CAGCTCCATGTTGGGCGCCATCTGCTGCGGAGCGCTCCCATGTCTTCTCTGCGGCAA 2421

Qy 320 rProLeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIl 340
Db 2422 GGCCATCTGGGTCAAGGGCTGGGGACACAC---CAGTATGGAGGCACTGGCGCGCTGAT 2478

Qy 340 eLeuLeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAlaAspAlaTy 360
Db 2479 CCTGAAAAGGTTGAGATCGCGCTCATCAACACGACACCTGC-----GAGAACCTCCT 2532

Qy 360 rGlnGlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspTh 380
Db 2533 GCGCAGCAGATCAGCGCGCATGATGTGCTGGGCTTCTCAGCGCGCGCTGGACTC 2592

Qy 380 rCysGlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisVal----- 398
Db 2593 CTGCCAGGTGATTCGGGGGACCCCTGTCACGCGTGGAGCGGATGGCGGATCTTCCA 2652

Qy 399 -ValGlyIleValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrTh 418
Db 2653 GCGCGGTGTGTGAGTGGGAGACGGCTGCGCTCAGAGGAACAGCAGCGCGGTGTACAC 2712

Qy 418 rLysValSerAlaTyrLeuAsnTrpIle 427
Db 2713 AAGGCTCCCTCTGTTTTCGGGACTGGATC 2740

RESULT 10
US-11-108-459-5
; Sequence 5, Application US/11108459
; GENERAL INFORMATION:
; APPLICANT: Zaas, Aimee
; APPLICANT: Schwartz, David A.
; APPLICANT: Peltz, Gary
; TITLE OF INVENTION: Polymorphic Plasminogen Genes and Uses Thereof
; FILE REFERENCE: DUKE-03785
; CURRENT APPLICATION NUMBER: US/11/108,459
; CURRENT FILING DATE: 2005-04-18
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5
; LENGTH: 2771
; TYPE: DNA
; ORGANISM: Mus musculus
US-11-108-459-5

Alignment Scores:
Pred. No.: 1.37e-39 Length: 2771
Score: 460.00 Matches: 121
Percent Similarity: 48.66% Conservative: 61
Best Local Similarity: 32.35% Mismatches: 127
Query Match: 19.64% Indels: 66
DB: 7 Gaps: 12

US-10-803-530-2 (1-435) x US-11-108-459-5 (1-2771)
Qy 76 AspGlyGluLeuAspCysProLeuGlyGluAspGluGluHisCysValLysSerPhePro 95
Db 1486 GACTCTGAGACAGACTGCTATGATGGGAATGGCAAGACTAT----- 1527
```



```

; TYPE: DNA
; ORGANISM: Homosapiens
US-60-732-162-277

Alignment Scores:
Pred. No.: 2,06e-27 Length: 4137
Score: 347.00 Matches: 122
Percent Similarity: 43.69% Conservative: 58
Best Local Similarity: 29.61% Mismatches: 138
Query Match: 14.82% Indels: 94
DB: 8 Gaps: 23

US-10-803-530-2 (1-435) x US-60-732-162-277 (1-4137)
QY 98 ProAlaValAlaValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAla 117
Db 1272 CCTCCCAAGCCCAAGTATTCTTCAAGAC-----CAAGTCTGCTGAGCTGT 1319
QY 118 ---ThrGlyAsnTrpPheSerAlaCysPheAspAsnThrGluAlaLeuAlaGluThr 136
Db 1320 GACACAGC-----TACAAAGTGCTGAAGGATAATGTGGAGATGACACATTCAGATT 1373
QY 137 AlaCysArgGlnMetGly---TyrSerSerLys---ProThrPheArgAlaValGluLeu 154
Db 1374 GAGTGTCTGAAGGATGGGACGTGGAGTAACAGATTCCACCTGTAAATAATTGTAGACTGT 1433
QY 155 GlyProAspGlnAspLeuAsp-----161
Db 1434 AGAGCCCCAGGAGAGCTGGACACACGGGCTGATCACCTTCTTACAAGGAACAACCTCAC 1493
QY 162 -----ValValGluLeuThrGluAsnSerGlnGlu-----LeuArgMetArgAsn 176
Db 1494 ACATACAGTCTGAGATCAATACTCTCTGTCAGGAGCCCTATTACAGATGCTCAACAT 1553
QY 177 SerSerGly-----ProCysLeuSerGly 184
Db 1554 AACACAGGTATATATACCTGTTCTGCCCAAGGAGCTGGATCAATAAAGTATTGGGAGA 1613
QY 185 SerLeuValSerLeuHisCysLeu---AlaCysGlySerLeuLysThr-----200
Db 1614 AGCCTACCCACG-----TGCCTTCAGAGTGTGGTCAAGCTCTCCCGCTCCCTGCCAAGC 1667
QY 201 -----ProArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnVal 218
Db 1668 CTGGTCAAGAGGATCATTTGGGGCGCAATGTGAGCCTGCTTCTCCCGTGGCAGGCC 1727
QY 219 SerIleGlnTyrAsp-----LysGlnHisValCysGlyGlySer 231
Db 1728 CTGATAGTGGTGGAGACACTTCGAGAGTCCCAATGACAGTGGTTTGGAGTGGGGCC 1787
QY 232 IleLeuAspProHisTrpValLeuThrAlaAlaHisCysPheArg-----246
Db 1788 CTGCTCTCTGCTGCTGGATCTCTCAGCAGCTCATGTGCTGCTGCCCTCCAGCGTAGAC 1847
QY 247 -----LysHisThrAspValPheAsnTrpLysValArgAla 258
Db 1848 ACCACGGTGATACCACTCTCAAGGAGCATGTCAACGCTCTAC---CTGGGCTTGCATGAT 1904
QY 259 GlySerAspLysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGlu 278
Db 1905 GTGCGAGACAAATCGGGGCGAGTCAACAGCTCAGCTGCGCGGAGTGTGCTCCACCCAGAC 1964
QY 279 PheAsnProMetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeu 298
Db 1965 TTCAAC---ATCCAAACTACCAACCATGATATAGCTCTGTGTCAGCTGCAGGAGCTGTG 2021
QY 299 ThrPheSerGlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrPro 318
Db 2022 CCCTGGGACCCACAGTTATGCTGTGCTGCTGCCAAGGCTTGAGCCTGAA---GGCCCG 2078
QY 319 AlaThrProLeu-----TrpIleIleGlyTrpGlyPheThrLysGlnAsn-----333
Db 2079 GCCCCCCACATGCTGGGCGCTGGTGGCGGCTGGGGCATCTCCAATCCCAATGTGACAGTG 2138

; Sequence 275, Application US/60732162
; GENERAL INFORMATION:
; APPLICANT: Belouchi, Abdelmajid
; APPLICANT: Raelson, John V
; APPLICANT: Bradley, Walter E
; APPLICANT: Paquin, Bruno
; APPLICANT: Fournier, Helene
; APPLICANT: Nguyen-Huu, Quynh
; APPLICANT: Croteau, Pascal
; APPLICANT: Allard, Rene
; APPLICANT: Debrus, Sophie
; APPLICANT: Berdewegh, Paul V
; APPLICANT: Little, Randall D
; APPLICANT: Keith, Tim
; APPLICANT: Segal, Jonathan
; TITLE OF INVENTION: Genemap of the Human Genes Associated With Asthma Disease
; FILE REFERENCE: 059908-5010-PR
; CURRENT APPLICATION NUMBER: US/60732,162
; CURRENT FILING DATE: 2003-11-02
; NUMBER OF SEQ ID NOS: 4417
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 275
; LENGTH: 4353
; TYPE: DNA
; ORGANISM: Homosapiens
US-60-732-162-275

Alignment Scores:
Pred. No.: 7,55e-25 Length: 4353
Score: 323.00 Matches: 109
Percent Similarity: 42.75% Conservative: 65
Best Local Similarity: 26.78% Mismatches: 133
Query Match: 13.79% Indels: 100
DB: 8 Gaps: 20

US-10-803-530-2 (1-435) x US-60-732-162-275 (1-4353)
QY 98 ProAlaValAlaValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAla 117
Db 1272 CCTCCCAAGCCCAAGTATTCTTCAAGAC-----CAAGTCTGCTGAGCTGT 1319
QY 118 ---ThrGlyAsnTrpPheSerAlaCysPheAspAsnThrGluAlaLeuAlaGluThr 136
Db 1320 GACACAGC-----TACAAAGTGCTGAAGGATAATGTGGAGATGACACATTCAGATT 1373
QY 137 AlaCysArgGlnMetGly---TyrSerSerLys---ProThrPheArgAlaValGluLeu 154
Db 1374 GAGTGTCTGAAGGATGGGACGTGGAGTAACAGATTCCACCTGTAAATAATTGTAGACTGT 1433
QY 155 GlyProAspGlnAspLeuAsp-----161
Db 1434 AGAGCCCCAGGAGAGCTGGACACACGGGCTGATCACCTTCTTACAAGGAACAACCTCAC 1493
QY 162 -----ValValGluLeuThrGluAsnSerGlnGlu-----LeuArgMetArgAsn 176
Db 1494 ACATACAGTCTGAGATCAATACTCTCTGTCAGGAGCCCTATTACAGATGCTCAACAT 1553
QY 177 SerSerGly-----ProCysLeuSerGly 184
Db 1554 AACACAGGTATATATACCTGTTCTGCCCAAGGAGCTGGATCAATAAAGTATTGGGAGA 1613
QY 185 SerLeuValSerLeuHisCysLeu---AlaCysGlySerLeuLysThr-----200
Db 1614 AGCCTACCCACG-----TGCCTTCAGAGTGTGGTCAAGCTCTCCCGCTCCCTGCCAAGC 1667
QY 201 -----ProArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnVal 218
Db 1668 CTGGTCAAGAGGATCATTTGGGGCGCAATGTGAGCCTGCTTCTCCCGTGGCAGGCC 1727
QY 219 SerIleGlnTyrAsp-----LysGlnHisValCysGlyGlySer 231
Db 1728 CTGATAGTGGTGGAGACACTTCGAGAGTCCCAATGACAGTGGTTTGGAGTGGGGCC 1787
QY 232 IleLeuAspProHisTrpValLeuThrAlaAlaHisCysPheArg-----246
Db 1788 CTGCTCTCTGCTGCTGGATCTCTCAGCAGCTCATGTGCTGCTGCCCTCCAGCGTAGAC 1847
QY 247 -----LysHisThrAspValPheAsnTrpLysValArgAla 258
Db 1848 ACCACGGTGATACCACTCTCAAGGAGCATGTCAACGCTCTAC---CTGGGCTTGCATGAT 1904
QY 259 GlySerAspLysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGlu 278
Db 1905 GTGCGAGACAAATCGGGGCGAGTCAACAGCTCAGCTGCGCGGAGTGTGCTCCACCCAGAC 1964
QY 279 PheAsnProMetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeu 298
Db 1965 TTCAAC---ATCCAAACTACCAACCATGATATAGCTCTGTGTCAGCTGCAGGAGCTGTG 2021
QY 299 ThrPheSerGlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrPro 318
Db 2022 CCCTGGGACCCACAGTTATGCTGTGCTGCTGCCAAGGCTTGAGCCTGAA---GGCCCG 2078
QY 319 AlaThrProLeu-----TrpIleIleGlyTrpGlyPheThrLysGlnAsn-----333
Db 2079 GCCCCCCACATGCTGGGCGCTGGTGGCGGCTGGGGCATCTCCAATCCCAATGTGACAGTG 2138

334 -----GlyGlyLys---MetSerAspIleLeuLeuGlnAlaSerVal 346
2139 GATGAGATCATCAGCAGTGGCACACGACCTTGTTCAGATGTCCTGTCAGATGATGTCAAGTTA 2198
QY 347 GlnValIleAspSerThrArgCysAsnAlaAspAlaTyrGlnGly-----362
2199 CCCGTGGTGCTCAGCTGAGTGAAGAACT-----AGCTATGAGTCCCGCTCGGCAAT 2252
QY 363 ---GluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys 381
2253 TACAGCGTCACGGAGAACATGTTCTGTGCTGCTACTACGAGGCGGCAAGACACACGTGC 2312
QY 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAsp-----GlnTrpHisVal 398
2313 CTTGAGATACGGTGGGCGCTTTGTCATCTTTGATGACTTGAGCCAGCGCTGGGTGTG 2372
QY 399 ValGlyIleValSerTrpGly-----TyrGlyCysGlyGlyProSerThrProGlyVal 416
2373 CAAGCGCTGGTGTCTCTGGGGGGACCTGAAGAAATGCGGCACAGCAGCTCTATGAGTGC 2432
QY 417 TyrThrLysValSerAlaTyrLeuAsnTrpIleTyr 428
2433 TACAAAGGCTTCAATTACGTGGACTGGGTGTGG 2468

RESULT 14
US-60-732-162-275
; Sequence 275, Application US/60732162
; GENERAL INFORMATION:
; APPLICANT: Belouchi, Abdelmajid
; APPLICANT: Raelson, John V
; APPLICANT: Bradley, Walter E
; APPLICANT: Paquin, Bruno
; APPLICANT: Fournier, Helene
; APPLICANT: Nguyen-Huu, Quynh
; APPLICANT: Croteau, Pascal
; APPLICANT: Allard, Rene
; APPLICANT: Debrus, Sophie
; APPLICANT: Berdewegh, Paul V
; APPLICANT: Little, Randall D
; APPLICANT: Keith, Tim
; APPLICANT: Segal, Jonathan
; TITLE OF INVENTION: Genemap of the Human Genes Associated With Asthma Disease
; FILE REFERENCE: 059908-5010-PR
; CURRENT APPLICATION NUMBER: US/60732,162
; CURRENT FILING DATE: 2003-11-02
; NUMBER OF SEQ ID NOS: 4417
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 275
; LENGTH: 4353
; TYPE: DNA
; ORGANISM: Homosapiens
US-60-732-162-275

Alignment Scores:
Pred. No.: 7,55e-25 Length: 4353
Score: 323.00 Matches: 109
Percent Similarity: 42.75% Conservative: 65
Best Local Similarity: 26.78% Mismatches: 133
Query Match: 13.79% Indels: 100
DB: 8 Gaps: 20

US-10-803-530-2 (1-435) x US-60-732-162-275 (1-4353)
QY 98 ProAlaValAlaValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAla 117
Db 1272 CCTCCCAAGCCCAAGTATTCTTCAAGAC-----CAAGTCTGCTGAGCTGT 1319
QY 118 ---ThrGlyAsnTrpPheSerAlaCysPheAspAsnThrGluAlaLeuAlaGluThr 136
Db 1320 GACACAGC-----TACAAAGTGCTGAAGGATAATGTGGAGATGACACATTCAGATT 1373
QY 137 AlaCysArgGlnMetGly---TyrSerSerLys---ProThrPheArgAlaValGluLeu 154
Db 1374 GAGTGTCTGAAGGATGGGACGTGGAGTAACAGATTCCACCTGTAAATAATTGTAGACTGT 1433
QY 155 GlyProAspGlnAspLeuAsp-----161
Db 1434 AGAGCCCCAGGAGAGCTGGACACACGGGCTGATCACCTTCTTACAAGGAACAACCTCAC 1493
QY 162 -----ValValGluLeuThrGluAsnSerGlnGlu-----LeuArgMetArgAsn 176
Db 1494 ACATACAGTCTGAGATCAATACTCTCTGTCAGGAGCCCTATTACAGATGCTCAACAT 1553
QY 177 SerSerGly-----ProCysLeuSerGly 184
Db 1554 AACACAGGTATATATACCTGTTCTGCCCAAGGAGCTGGATCAATAAAGTATTGGGAGA 1613
QY 185 SerLeuValSerLeuHisCysLeu---AlaCysGlySerLeuLysThr-----200
Db 1614 AGCCTACCCACG-----TGCCTTCAGAGTGTGGTCAAGCTCTCCCGCTCCCTGCCAAGC 1667
QY 201 -----ProArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnVal 218
Db 1668 CTGGTCAAGAGGATCATTTGGGGCGCAATGTGAGCCTGCTTCTCCCGTGGCAGGCC 1727
QY 219 SerIleGlnTyrAsp-----LysGlnHisValCysGlyGlySer 231
Db 1728 CTGATAGTGGTGGAGACACTTCGAGAGTCCCAATGACAGTGGTTTGGAGTGGGGCC 1787
QY 232 IleLeuAspProHisTrpValLeuThrAlaAlaHisCysPheArg-----246
Db 1788 CTGCTCTCTGCTGCTGGATCTCTCAGCAGCTCATGTGCTGCTGCCCTCCAGCGTAGAC 1847
QY 247 -----LysHisThrAspValPheAsnTrpLysValArgAla 258
Db 1848 ACCACGGTGATACCACTCTCAAGGAGCATGTCAACGCTCTAC---CTGGGCTTGCATGAT 1904
QY 259 GlySerAspLysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGlu 278
Db 1905 GTGCGAGACAAATCGGGGCGAGTCAACAGCTCAGCTGCGCGGAGTGTGCTCCACCCAGAC 1964
QY 279 PheAsnProMetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeu 298
Db 1965 TTCAAC---ATCCAAACTACCAACCATGATATAGCTCTGTGTCAGCTGCAGGAGCTGTG 2021
QY 299 ThrPheSerGlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrPro 318
Db 2022 CCCTGGGACCCACAGTTATGCTGTGCTGCTGCCAAGGCTTGAGCCTGAA---GGCCCG 2078
QY 319 AlaThrProLeu-----TrpIleIleGlyTrpGlyPheThrLysGlnAsn-----333
Db 2079 GCCCCCCACATGCTGGGCGCTGGTGGCGGCTGGGGCATCTCCAATCCCAATGTGACAGTG 2138
```



```

Db 1819 CCTGAAAGACAGTTGCAGTGTATTATGGCTGGGGCTACACT-----GGATTG 1866
Qy 337 MetSer-----AspIleLeuGlnAlaSerValGlnValIleAspSerThrArgCys 354
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 1867 ATCAACTATGATGGCTATTACGAGTGCACATCTCTATATAATGGGAAATGAGAAATGC 1926
Qy 355 AsnAlaAspAspAlaTyrGlnGluValThr-----GluLysMetMetCysAlaGly 372
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 1927 AGC-----CAGCATCATCGAGGGAAGGTGACTCTGAATGAGTCTGAAATATGTGCTGGG 1980
Qy 373 IleProGluGlyGlyValAspThrCysGlnGlyAspSerGlyGlyProLeuMetTyrGln 392
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 1981 GCTGAAAGATTGGATCAGGACCATGTGAGGGGGAATTATGGTGGCCCACTGTGTTGTGAG 2040
Qy 393 SerAspGlnTrpHis---ValValGlyIleValSerTrpGlyTyrGlyCysGlyGlyPro 411
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 2041 CAACATAAAATGAGAAATGGTTCTTGGTGTTCATTGTTCTCTGGTGGTGGATGGCCATTCCA 2100
Qy 412 SerThrProGlyValTyrThrLysValSerAlaTyrLeuAsnTrpIleTyrAsnVal 430
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 2101 AATCGTCCTGGTATTTTGTCCGAGTAGCATATTATGCAAAATGGATACACAAAATT 2157

```

Search completed: December 7, 2005, 00:57:52
Job time : 72 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2005, 13:49:20 ; Search time 12 Seconds
(without alignments)
208.093 Million cell updates/sec

Title: US-10-803-530-2

Perfect score: 2342

Sequence: 1 MDPDSQPLNSLDVKPLRPK.....VYTKVSAYLNWYNWKAEL 435

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 41597 seqs, 5740495 residues

Total number of hits satisfying chosen parameters: 41597

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA New:*

- 1: /cgn2_6/ptodata/2/paa/pct NEW COMB.pcp.*
- 2: /cgn2_6/ptodata/2/paa/US06 NEW COMB.pcp.*
- 3: /cgn2_6/ptodata/2/paa/US07 NEW COMB.pcp.*
- 4: /cgn2_6/ptodata/2/paa/US08 NEW COMB.pcp.*
- 5: /cgn2_6/ptodata/2/paa/US09 NEW COMB.pcp.*
- 6: /cgn2_6/ptodata/2/paa/US10 NEW COMB.pcp.*
- 7: /cgn2_6/ptodata/2/paa/US11 NEW COMB.pcp.*
- 8: /cgn2_6/ptodata/2/paa/US60 NEW COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	2342	100.0	461	US-10-556-478-3
2	2333	99.6	435	US-11-268-554-221
3	2328	99.4	437	US-11-268-554-220
4	2297.5	98.1	432	US-10-206-921A-330
5	2096	89.5	391	US-10-556-478-2
6	1296	55.3	486	US-10-556-478-1
7	658	28.1	453	US-10-206-921A-64
8	574	24.5	457	US-60-732-162-488
9	478	20.4	314	US-10-868-184C-4975
10	478	20.4	340	US-10-556-478-5
11	475	20.3	423	US-10-206-921A-320
12	475	20.3	423	US-11-265-762-146
13	467.5	20.0	290	US-11-264-096-1294
14	463	19.8	855	US-11-268-554-284
15	459.5	19.6	812	US-11-108-459-6
16	454.5	19.4	287	US-10-556-478-4
17	418.5	17.9	798	US-11-108-459-8
18	356	15.2	431	PCT-US05-39108-1
19	355.5	15.2	415	US-11-166-028-1
20	355.5	15.2	461	US-11-172-459-1
21	347	14.8	728	US-60-732-162-278
22	343	14.6	261	PCT-US05-4045-12
23	323	13.8	699	US-60-732-162-276
24	299.5	12.8	723	US-60-732-162-1604
25	299.5	12.8	728	US-10-555-925-1

ALIGNMENTS

RESULT 1

US-10-556-478-3
; Sequence 3, Application US/10556478

; GENERAL INFORMATION:

; APPLICANT: diabex, Inc.

; APPLICANT: Pilkington, Glenn

; APPLICANT: Keller, Gilbert-Andre

; APPLICANT: Li, Wenlu

; APPLICANT: Corral, Laura

; APPLICANT: Simon, Iris

; TITLE OF INVENTION: Ovr115 Antibody Compositions and Methods of Use

; FILE REFERENCE: DEX-0484

; CURRENT APPLICATION NUMBER: US/10/556,478

; CURRENT FILING DATE: 2005-11-14

; PRIOR APPLICATION NUMBER: US 60/559,730

; PRIOR FILING DATE: 2004-04-05

; PRIOR APPLICATION NUMBER: US 60/471,068

; PRIOR FILING DATE: 2003-05-16

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 3

; LENGTH: 461

; TYPE: PRT

; ORGANISM: Artificial sequence

; FEATURE:

; OTHER INFORMATION: Synthetic

US-10-556-478-3

Query Match 100.0% Score 2342; DB 6; Length 461;

Best Local Similarity 100.0%; Pred. No. 2.1e-248; Indels 0; Gaps 0;
Matches 435; Conservative 0; Mismatches 0;

QY	1	MDPDSQPLNSLDVKPLRPKPRIPMETFRKVGIPPIIIALLSLASIIIVVLIKVLIDKYYF	60
DB	1	MDPDSQPLNSLDVKPLRPKPRIPMETFRKVGIPPIIIALLSLASIIIVVLIKVLIDKYYF	60
QY	61	LCGQPLHFIPRKQLCDGELDCPLGDEDEHCVKSPFEGPAVAVRLSKDRSTLQVLSATGN	120
DB	61	LCGQPLHFIPRKQLCDGELDCPLGDEDEHCVKSPFEGPAVAVRLSKDRSTLQVLSATGN	120
QY	121	WFSACDFNFTALACCKSLKTRPVVGBEASVDSWPQVSIQYDKHVCVGSILDPHWLIT	240
DB	121	WFSACDFNFTALACCKSLKTRPVVGBEASVDSWPQVSIQYDKHVCVGSILDPHWLIT	240
QY	181	CLUGSLVSLHCLACCKSLKTRPVVGBEASVDSWPQVSIQYDKHVCVGSILDPHWLIT	240
DB	181	CLUGSLVSLHCLACCKSLKTRPVVGBEASVDSWPQVSIQYDKHVCVGSILDPHWLIT	240

```
QY 241 AAHCFRKHDTVFNWVKVRAAGSKLGSFPLSLAVAKIIIEFNPMYPKNDIALMKLQPLTF 300
Db 241 AAHCFRKHDTVFNWVKVRAAGSKLGSFPLSLAVAKIIIEFNPMYPKNDIALMKLQPLTF 300
QY 301 SGTVRPCLPFFDEELTPATLWIIGWFTKQNGKMSDILLQASVQVIDSTRCNADDAY 360
Db 301 SGTVRPCLPFFDEELTPATLWIIGWFTKQNGKMSDILLQASVQVIDSTRCNADDAY 360
QY 361 QGEVTEKMCAGIPEGGVDTCCGDSGGLMYQSDQHWVGVISWGYCGGPGSTPGVYTKV 420
Db 361 QGEVTEKMCAGIPEGGVDTCCGDSGGLMYQSDQHWVGVISWGYCGGPGSTPGVYTKV 420
QY 421 SAYLNWIYNWKAEL 435
Db 421 SAYLNWIYNWKAEL 435
```

RESULT 2

```
US-11-268-554-221
; Sequence 221, Application US/11268554
; GENERAL INFORMATION:
; APPLICANT: DOMON, Bruno et al.
; TITLE OF INVENTION: PANCREATIC CANCER TARGETS AND USES
; FILE OF INVENTION: THEREOF
; FILE REFERENCE: CL001557
; CURRENT APPLICATION NUMBER: US/11/268,554
; CURRENT FILING DATE: 2005-11-08
; NUMBER OF SEQ ID NOS: 1004
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 221
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-268-554-221
```

```
Query Match 99.6%; Score 2333; DB 7; Length 435;
Best Local Similarity 99.8%; Pred. No. 1.9e-247;
Matches 434; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 MDPSDQPLNSLDVVKPLRKPRIPMETFRKVGIPPIIIALLSLASIIIVVLIKVLDKYFL 60
Db 1 MDPSDQPLNSLDVVKPLRKPRIPMETFRKVGIPPIIIALLSLASIIIVVLIKVLDKYFL 60
QY 61 LCGQPLHPIPRKQLCDGELDCPLGEDEBHCVKSPPEGPAVAVRLSKORSTLQVLDSATGN 120
Db 61 LCGQPLHPIPRKQLCDGELDCPLGEDEBHCVKSPPEGPAVAVRLSKORSTLQVLDSATGN 120
QY 121 WFSACFDNFTALAEACRQMGYSKPTFRAVEIGPDQDLVDVEITENSQELMRNSSGP 180
Db 121 WFSACFDNFTALAEACRQMGYSKPTFRAVEIGPDQDLVDVEITENSQELMRNSSGP 180
QY 181 CLSGSLVSLHCLACGSKLTPRVVVGEEASVDSWPMQVSIQYDKQHVCGGSILDPHWLTA 240
Db 181 CLSGSLVSLHCLACGSKLTPRVVVGEEASVDSWPMQVSIQYDKQHVCGGSILDPHWLTA 240
QY 241 AAHCFRKHDTVFNWVKVRAAGSKLGSFPLSLAVAKIIIEFNPMYPKNDIALMKLQPLTF 300
Db 241 AAHCFRKHDTVFNWVKVRAAGSKLGSFPLSLAVAKIIIEFNPMYPKNDIALMKLQPLTF 300
QY 301 SGTVRPCLPFFDEELTPATLWIIGWFTKQNGKMSDILLQASVQVIDSTRCNADDAY 360
Db 301 SGTVRPCLPFFDEELTPATLWIIGWFTKQNGKMSDILLQASVQVIDSTRCNADDAY 360
QY 361 QGEVTEKMCAGIPEGGVDTCCGDSGGLMYQSDQHWVGVISWGYCGGPGSTPGVYTKV 420
Db 361 QGEVTEKMCAGIPEGGVDTCCGDSGGLMYQSDQHWVGVISWGYCGGPGSTPGVYTKV 420
QY 421 SAYLNWIYNWKAEL 435
Db 421 SAYLNWIYNWKAEL 435
```

RESULT 3

```
US-11-268-554-220
; Sequence 220, Application US/11268554
; GENERAL INFORMATION:
; APPLICANT: DOMON, Bruno et al.
; TITLE OF INVENTION: PANCREATIC CANCER TARGETS AND USES
; FILE OF INVENTION: THEREOF
; FILE REFERENCE: CL001557
; CURRENT APPLICATION NUMBER: US/11/268,554
; CURRENT FILING DATE: 2005-11-08
; NUMBER OF SEQ ID NOS: 1004
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 220
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-268-554-220
```

```
Query Match 99.4%; Score 2328; DB 7; Length 437;
Best Local Similarity 99.8%; Pred. No. 6.6e-247;
Matches 433; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 DPDSQPLNSLDVVKPLRKPRIPMETFRKVGIPPIIIALLSLASIIIVVLIKVLDKYFL 61
Db 4 DPDSQPLNSLDVVKPLRKPRIPMETFRKVGIPPIIIALLSLASIIIVVLIKVLDKYFL 63
QY 62 CGQPLHPIPRKQLCDGELDCPLGEDEBHCVKSPPEGPAVAVRLSKORSTLQVLDSATGN 121
Db 64 CGQPLHPIPRKQLCDGELDCPLGEDEBHCVKSPPEGPAVAVRLSKORSTLQVLDSATGN 123
QY 122 FSACFDNFTALAEACRQMGYSKPTFRAVEIGPDQDLVDVEITENSQELMRNSSGP 181
Db 124 FSACFDNFTALAEACRQMGYSKPTFRAVEIGPDQDLVDVEITENSQELMRNSSGP 183
QY 182 LSGSLVSLHCLACGSKLTPRVVVGEEASVDSWPMQVSIQYDKQHVCGGSILDPHWLTA 241
Db 184 LSGSLVSLHCLACGSKLTPRVVVGEEASVDSWPMQVSIQYDKQHVCGGSILDPHWLTA 243
QY 242 AHCFRKHDTVFNWVKVRAAGSKLGSFPLSLAVAKIIIEFNPMYPKNDIALMKLQPLTF 301
Db 244 AHCFRKHDTVFNWVKVRAAGSKLGSFPLSLAVAKIIIEFNPMYPKNDIALMKLQPLTF 303
QY 302 GTVRPCLPFFDEELTPATLWIIGWFTKQNGKMSDILLQASVQVIDSTRCNADDAY 361
Db 304 GTVRPCLPFFDEELTPATLWIIGWFTKQNGKMSDILLQASVQVIDSTRCNADDAY 363
QY 362 GEVTEKMCAGIPEGGVDTCCGDSGGLMYQSDQHWVGVISWGYCGGPGSTPGVYTKV 421
Db 364 GEVTEKMCAGIPEGGVDTCCGDSGGLMYQSDQHWVGVISWGYCGGPGSTPGVYTKV 423
QY 422 AYLNWIYNWKAEL 435
Db 424 AYLNWIYNWKAEL 437
```

RESULT 4

```
US-10-206-921A-330
; Sequence 330, Application US/10206921A
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: 39780-3430R1C515
; CURRENT APPLICATION NUMBER: US/10/206,921A
```



```
; CURRENT FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: US 10/052,586
; FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: PCT/US01/06520
; FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/232,887
; FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 330
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-206-921A-330

Query Match      98.1%; Score 2297.5; DB 6; Length 432;
Best Local Similarity 98.8%; Pred. No. 1.4e-243;
Matches 429; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY  2 DPDSQPLNSLDVKPLRPMPMETFRKVGIPPIIIALLSLASIIIVVILKVIDKYFL 61
Db  4 DPDSQPLNSLDVKPLRPMPMETFRKVGIPPIIIALLSLASIIIVVILKVIDKYFL 63

QY  62 CGQPLHFIIPRKQCDGELDCPLGEDEEHCVKSPFEGPAVAVRLSKDRSTLQVLD SATGNW 121
Db  64 CGQPLHFIIPRKQCDGELDCPLGEDEEHCVKSPFEGPAVAVRLSKDRSTLQVLD SATGNW 123

QY  122 FSACFDNFTALAEACRQMGYSKPTFRAVEIGPDQLDVVEITENSQELMRNSSGPC 181
Db  124 FSACFDNFTALAEACRQMGYS-----RAVEIGPDQLDVVEITENSQELMRNSSGPC 178

QY  182 LSGSLVSLHCLACGKSLKTPRVVGGEEASVDSWPQVSIQYDKQHVCGVILDPHVLTA 241
Db  179 LSGSLVSLHCLACGKSLKTPRVVGGEEASVDSWPQVSIQYDKQHVCGVILDPHVLTA 238

QY  242 AHCPRKHTDVFNKVRAGSKLGSPPSLAVAKIIIEFNPMYPKNDIALMKLQPLTFS 301
Db  239 AHCPRKHTDVFNKVRAGSKLGSPPSLAVAKIIIEFNPMYPKNDIALMKLQPLTFS 298

QY  302 GTVRPICLPFFDEELTPATPLIIGWFTKQNGKMSDILLQASVQVIDSTRCNADDAYQ 361
Db  299 GTVRPICLPFFDEELTPATPLIIGWFTKQNGKMSDILLQASVQVIDSTRCNADDAYQ 358

QY  362 GVETKMMKAGIPEGGVDTCCQDSGGPLMYQSDQHVHVGIVSWGCGGSPTPGVYTKVS 421
Db  359 GVETKMMKAGIPEGGVDTCCQDSGGPLMYQSDQHVHVGIVSWGCGGSPTPGVYTKVS 418

QY  422 AYLNWIYNWKAEL 435
Db  419 AYLNWIYNWKAEL 432

RESULT 5
US-10-556-478-2
; Sequence 2, Application US/10556478
; GENERAL INFORMATION:
; APPLICANT: diadexus, Inc.
; APPLICANT: Pilkington, Glenn
; APPLICANT: Keller, Gilbert-Andre
; APPLICANT: Li, Wenlu
; APPLICANT: Corral, Laura
; APPLICANT: Simon, Iris
; TITLE OF INVENTION: Ovr115 Antibody Compositions and Methods of Use
; FILE REFERENCE: DEX-0484
; CURRENT APPLICATION NUMBER: US/10/556,478
; PRIOR FILING DATE: 2005-11-14
; PRIOR APPLICATION NUMBER: US 60/559,730
; PRIOR FILING DATE: 2004-04-05
; PRIOR APPLICATION NUMBER: US 60/471,068
; PRIOR FILING DATE: 2003-05-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 391

; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-556-478-2

Query Match      89.5%; Score 2096; DB 6; Length 391;
Best Local Similarity 99.7%; Pred. No. 1.4e-221;
Matches 384; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY  51 IKVILDKYFLCGQPLHFIIPRKQCDGELDCPLGEDEEHCVKSPFEGPAVAVRLSKDRST 110
Db  1 MKVILDKYFLCGQPLHFIIPRKQCDGELDCPLGEDEEHCVKSPFEGPAVAVRLSKDRST 60

QY  111 LQVLD SATGNWFSACFDNFTALAEACRQMGYSKPTFRAVEIGPDQLDVVEITENSQ 170
Db  61 LQVLD SATGNWFSACFDNFTALAEACRQMGYSKPTFRAVEIGPDQLDVVEITENSQ 120

QY  171 ELRNRSSGPCLSGLSVSLHCLACGKSLKTPRVVGGEEASVDSWPQVSIQYDKQHVCGG 230
Db  121 ELRNRSSGPCLSGLSVSLHCLACGKSLKTPRVVGGEEASVDSWPQVSIQYDKQHVCGG 180

QY  231 SILDPHWLVLTAAHCFRKHDTVFNKVRAGSKLGSPPSLAVAKIIIEFNPMYPKNDIA 290
Db  181 SILDPHWLVLTAAHCFRKHDTVFNKVRAGSKLGSPPSLAVAKIIIEFNPMYPKNDIA 240

QY  291 LMKLQPLTFSGTVRPICLPFFDEELTPATPLIIGWFTKQNGKMSDILLQASVQVID 350
Db  241 LMKLQPLTFSGTVRPICLPFFDEELTPATPLIIGWFTKQNGKMSDILLQASVQVID 300

QY  351 STRCNADDAYQGVETKMMKAGIPEGGVDTCCQDSGGPLMYQSDQHVHVGIVSWGCGG 410
Db  301 STRCNADDAYQGVETKMMKAGIPEGGVDTCCQDSGGPLMYQSDQHVHVGIVSWGCGG 360

QY  411 PSTFGVYTKVSAYLNWIYNWKAEL 435
Db  361 PSTFGVYTKVSAYLNWIYNWKAEL 385

RESULT 6
US-10-556-478-1
; Sequence 1, Application US/10556478
; GENERAL INFORMATION:
; APPLICANT: diadexus, Inc.
; APPLICANT: Pilkington, Glenn
; APPLICANT: Keller, Gilbert-Andre
; APPLICANT: Li, Wenlu
; APPLICANT: Corral, Laura
; APPLICANT: Simon, Iris
; TITLE OF INVENTION: Ovr115 Antibody Compositions and Methods of Use
; FILE REFERENCE: DEX-0484
; CURRENT APPLICATION NUMBER: US/10/556,478
; CURRENT FILING DATE: 2005-11-14
; PRIOR APPLICATION NUMBER: US 60/559,730
; PRIOR FILING DATE: 2004-04-05
; PRIOR APPLICATION NUMBER: US 60/471,068
; PRIOR FILING DATE: 2003-05-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 486
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-556-478-1

Query Match      55.3%; Score 1296; DB 6; Length 486;
Best Local Similarity 60.1%; Pred. No. 7.9e-134;
Matches 270; Conservative 21; Mismatches 56; Indels 102; Gaps 9;

QY  48 VVILKVIDKYFLCGQPLHFIIPRKQCDGELDCPLGEDEEHCVKSPFEGPAVAVR---- 103
Db  69 MAIRYIADKHNMLGG-----CP-----KERAELSMLEGAVIDIRYGV 107
```

QY 104 ---LSKDRSTLOV-----LDS 116
Db 108 RIAYSKDFETLKVDFLSKLPMLKMFEDRLCHKTYLNGHDVTHPDRLYDALDVLYMDP 167
QY 117 ATGNWFS--ACFNFTAL--ASTACROMGYSKP-----TPRAVEIGDQDOLDVVEIT 166
Db 168 MCLDAPFKLVCFKKRIEAPQIDIKLKSSKIYIAIPLQGWQATFGGDHFPKSDL-----VP 223
QY 167 ENSQELMRNMSGPCLSGSLVSLHCLACGKSLKTPRVVGEEASVDSWPQVSIQYDKOH 226
Db 224 RHNQ-----TSLYKAGFENLYFQGVVGEEASVDSWPQVSIQYDKOH 267
QY 227 VCGGSLDPHVLVTAHCFRKHDTDFNWKVRAGSKLGFPSLAVAKIIIIIEFNPMYKPD 286
Db 268 VCGGSLDPHVLVTAHCFRKHDTDFNWKVRAGSKLGFPSLAVAKIIIIIEFNPMYKPD 327
QY 287 NDIALMKLOPPLTFSGTVPICLPFFDELTTPATPLWIIIGWFTKQNGKMSDILLQASV 346
Db 328 NDIALMKLOPPLTFSGTVPICLPFFDELTTPATPLWIIIGWFTKQNGKMSDILLQASV 387
QY 347 QVIDSTRCNADDAAYQGEVTEKMCAGIPGGVDTCQDGGGGLMYQSDQVHVGIVSWGY 406
Db 388 QVIDSTRCNADDAAYQGEVTEKMCAGIPGGVDTCQDGGGGLMYQSDQVHVGIVSWGY 447
QY 407 GCGGPTPGVYTKVSAYLWNIYVNWKAEL 435
Db 448 GCGGPTPGVYTKVSAYLWNIYVNWKAEL 476

RESULT 7
US-10-206-921A-64
; Sequence 64, Application US/10206921A
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: 39780-3430R1C515
; CURRENT APPLICATION NUMBER: US/10/206,921A
; CURRENT FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: US 10/052,586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: PCT/US01/06520
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/232,887
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 64
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-206-921A-64

Query Match 28.1%; Score 658; DB 6; Length 453;
Best Local Similarity 35.7%; Pred. No. 5.1e-64;
Matches 158; Conservative 75; Mismatches 146; Indels 64; Gaps 18;
QY 22 IPMTFRKVGPIPIIALLSLASIIIVVLKIVLD---KY-----YFLCGQPLHFIPRKQL 74
Db 44 LPLKFFPIIIGITIALILALA-----IGLGIHFDCKGKRCRSSFKC---IELIAR--- 91
QY 75 CDGELDCPLGEDEBHCVKSPFEGPAVAVRLSKDRSTLOVLDATGNWFSACFNFTAL 134
Db 92 CDGVSDCKGDEDEYRC-----VRGGQNAVLOVFTAA--SWKTMCSDDMKGHYA 138

QY 135 ETACROMGYSS-----KPTTFAVEIGDQDOLDVVEITENSQELMRNMSGPCL 182
Db 139 NVACAQLGFPFVYSSDNLRVSSLEGQFREBFVSIIDHLLPDDKVTAALHHSVVVREG---CA 195
QY 183 SGLSVSLHCLACG-KSLKTPRVVGEEASVDSWPQVSIQYDKOHVCGGSLDPHVLVTA 241
Db 196 SGHVVTILOCTACGHRRGYSRIVGGNMSLISQWPMQASLOFQGYHLGCGSVITPLWITA 255
QY 242 AHCFRKHDTVF--NWKVRAGSKL--GSPFSLAVAKIIIIIEFNPMY-PK--DNDIALMK 293
Db 256 AHCV---YDLYLPKSWIOVGLVSLDNDPAPSHLVEKIV---YHSKYKPKRLGNDIALMK 309
QY 294 LQPLTFSGTVPICLPFFDELTTPATPLWIIIGWFTKQNGKMSDILLQASVQVIDSTR 353
Db 310 LAGPLTFNEMIQVCLPNSSENFDPGKVCWCTSGWGAT-EDGGDASPVLNHAAVPLINSKI 368
QY 354 CNADDAAYQGEVTEKMCAGIPGGVDTCQDGGGGLMYQSDQ-VHVVGVISWVGCGGSPS 412
Db 369 CNHRDVTGGIISPSMLCAGYLTGGVDSQDGGGGLVQCERRLWKLVGATSGFGTCAEVN 428
QY 413 TPGVYTKVSAYLWNIYVNWKAEL 435
Db 429 KPGVYTRVTSFLDNIHQMERDL 451

RESULT 8
US-60-732-162-488
; Sequence 488, Application US/60732162
; GENERAL INFORMATION:
; APPLICANT: Belouchi, Abdelmajid
; APPLICANT: Raelson, John V
; APPLICANT: Bradley, Walter E
; APPLICANT: Paquin, Bruno
; APPLICANT: Fournier, Helene
; APPLICANT: Nguyen-Huu, Quynh
; APPLICANT: Croteau, Pascal
; APPLICANT: Allard, Rene
; APPLICANT: Debrus, Sophie
; APPLICANT: Berdewegh, Paul V
; APPLICANT: Little, Randall D
; APPLICANT: Keith, Tim
; APPLICANT: Segal, Jonathan
; TITLE OF INVENTION: Genemap of the Human Genes Associated With Asthma Disease
; FILE REFERENCE: 059908-5010-PR
; CURRENT APPLICATION NUMBER: US/60/732,162
; CURRENT FILING DATE: 2005-11-02
; NUMBER OF SEQ ID NOS: 4417
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 488
; LENGTH: 457
; TYPE: PRT
; ORGANISM: Homosapiens
US-60-732-162-488

Query Match 24.5%; Score 574; DB 8; Length 457;
Best Local Similarity 32.2%; Pred. No. 8.2e-55;
Matches 129; Conservative 70; Mismatches 151; Indels 50; Gaps 12;
QY 60 FLC---GQPLHFIPRKQLCDGE--LDCPLGEDEBHCVKSPFEGPAVAVRLSKDRSTLOVL 114
Db 71 YLCPAASQFI-----SGTLQDEEITLSCSEASAEALLPALPK--TVSFRINSEDFLLBAQ 124
QY 115 DSATGNWFSACFNFTALAEATACRQMGYSKPTPRAVEIGDQDOLDVVEITENSQE--- 171
Db 125 VRDQPRLLLVCHEGWSPALGLQICWSLGH-----LRLTHKGVNLTDKLNSQSEFA 176
QY 172 -----LRMRNMSGPCLSGSLVSLHCLACGKSLKTPRVVGEEASVDSWPQV 218
Db 177 QLSPLRGLFBEAWQPRNN---CTSGQVSLRSCGCGARPLASRIVGGQSVAPGRWPQA 233
QY 219 STQYDKOHVCGGSLDPHVLVTAHCFR--KHTDVFNWKVPAGSKLGSFPLAVAKIIII 276

Db 234 SVALGFRHTRCGSVLAPRWVTAACHMHSFRLARLSSWRVHAG---LVSHSAVRPHQCAL 290
Qy 277 IEF---NPMYKDN---DIALMKLQPLTFSTGTVRPICLPFFDEBELTPTATPLWIIGWGT 330
Db 291 VERIIPHPYSAQNHDYDVALRLQALNFSDTVGAVCLPAKEQHFPKGRSCWVGWGT 350
Qy 331 KONGKMSDILLQASVQVIDSTRCNADAYQGEVTEKMKCAGIPGGGVDTCCGDSGGPLM 390
Db 351 HPSHTYSMDLQDTVPVLFSTQLCNSSCVSYGALTFRMLCAGYLDGRADACQCGSGGPLV 410
Qy 391 Y-QSDWHVGVSVSGCGGPSTPGVYTKVSAYLNWIYN 429
Db 411 CPDGTWRLGVVSWGRACABPHNPGVYAKVAEFLDWIHD 450

RESULT 9

US-10-868-184C-4975
; Sequence 4975, Application US/10868184C
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS805
; CURRENT APPLICATION NUMBER: US/10/868,184C
; PRIOR FILING DATE: 2004-06-16
; PRIOR APPLICATION NUMBER: 60/278,650
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 09/833,245
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: PCT/US01/11988
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: PCT/US00/06043
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06012
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06058
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06044
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06059
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06042
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06014
; PRIOR FILING DATE: 2000-03-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 13046
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4975
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-868-184C-4975

Query Match 20.4%; Score 478; DB 6; Length 314;
Best Local Similarity 37.8%; Pred. No. 1.5e-44;
Matches 101; Conservative 48; Mismatches 86; Indels 32; Gaps 8;

Qy 178 SGPCLSGSLVSLHCLACGKSLKTPRVVGGEASVDSWPQVSIQYDKQHVCGSGILDHPW 237
Db 30 SGP-----CGRRVITSRIVGGEDELGRWPMQGSRLRLWDSHVCGVSLLSHRW 76
Qy 238 VLTAACHCRKHTDVFN---WKVRAGSDKLGSPFSLAVAK-----IIIEFNPMYPKDN 287
Db 77 ALTAACHCFETYDLSLSDPSGMMVQFG--QLTSMPSFWSLQAYYTRYFVSNVILSPRYLGS 134
Qy 288 --DIALMKLQPLTFSTGTVRPICLPFFDEBELTPTATPLWIIGWFTKQNGKMSDILLQ-A 344
Db 135 PYDIALVKLSAPVTVTKHIQICLQASTFEFENRTDCWVTGWGVIKDEALPSPHTLQEV 194
Qy 345 SVQVIDSTRCN---ADDAVQGEVTEKMKCAGIPGGGVDTCCGDSGGPL-MYQSDQWHVVG 400
Db 195 QVAIIINSMCNHLFLKYSFRKDIIFDMVCAGNAQGGKDACFGDSGGPLACNKNGLWYQIG 254

Qy 401 IVSHGYCGGPGSTPGVYTKVSAYLNWI 427
Db 255 VWSMGVCGGRNRPFGVYTNISHHFEWI 281

RESULT 10

US-10-556-478-5
; Sequence 5, Application US/10556478
; GENERAL INFORMATION:
; APPLICANT: diaDexus, Inc.
; APPLICANT: Pilkington, Glenn
; APPLICANT: Keller, Gilbert-Andre
; APPLICANT: Li, Wenlu
; APPLICANT: Corral, Laura
; APPLICANT: Simon, Iris
; TITLE OF INVENTION: Ovr115 Antibody Compositions and Methods of Use
; FILE REFERENCE: DEX-0484
; CURRENT APPLICATION NUMBER: US/10/556,478
; PRIOR FILING DATE: 2005-11-14
; PRIOR APPLICATION NUMBER: US 60/559,730
; PRIOR FILING DATE: 2004-04-05
; PRIOR APPLICATION NUMBER: US 60/471,068
; PRIOR FILING DATE: 2003-05-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 340
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-556-478-5

Query Match 20.4%; Score 478; DB 6; Length 340;
Best Local Similarity 37.8%; Pred. No. 1.7e-44;
Matches 101; Conservative 48; Mismatches 86; Indels 32; Gaps 8;

Qy 178 SGPCLSGSLVSLHCLACGKSLKTPRVVGGEASVDSWPQVSIQYDKQHVCGSGILDHPW 237
Db 30 SGP-----CGRRVITSRIVGGEDELGRWPMQGSRLRLWDSHVCGVSLLSHRW 76
Qy 238 VLTAACHCRKHTDVFN---WKVRAGSDKLGSPFSLAVAK-----IIIEFNPMYPKDN 287
Db 77 ALTAACHCFETYDLSLSDPSGMMVQFG--QLTSMPSFWSLQAYYTRYFVSNVILSPRYLGS 134
Qy 288 --DIALMKLQPLTFSTGTVRPICLPFFDEBELTPTATPLWIIGWFTKQNGKMSDILLQ-A 344
Db 135 PYDIALVKLSAPVTVTKHIQICLQASTFEFENRTDCWVTGWGVIKDEALPSPHTLQEV 194
Qy 345 SVQVIDSTRCN---ADDAVQGEVTEKMKCAGIPGGGVDTCCGDSGGPL-MYQSDQWHVVG 400
Db 195 QVAIIINSMCNHLFLKYSFRKDIIFDMVCAGNAQGGKDACFGDSGGPLACNKNGLWYQIG 254
Qy 401 IVSHGYCGGPGSTPGVYTKVSAYLNWI 427
Db 255 VWSMGVCGGRNRPFGVYTNISHHFEWI 281

RESULT 11

US-10-206-921A-320
; Sequence 320, Application US/10206921A
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin

```
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ FILE REFERENCE: 39780-3430R1C515
/ CURRENT APPLICATION NUMBER: US/10/206,921A
/ CURRENT FILING DATE: 2002-07-26
/ PRIOR APPLICATION NUMBER: US 10/052,586
/ PRIOR FILING DATE: 2002-01-15
/ PRIOR APPLICATION NUMBER: PCT/US01/06520
/ PRIOR FILING DATE: 2001-02-28
/ PRIOR APPLICATION NUMBER: US 60/232,887
/ PRIOR FILING DATE: 2000-09-15
/ NUMBER OF SEQ ID NOS: 612
/ SEQ ID NO 320
/ LENGTH: 423
/ TYPE: PRT
/ ORGANISM: Homo Sapien
/ ORGANISM: Homo Sapien
US-10-206-921A-320

Query Match      20.3%; Score 475; DB 6; Length 423;
Best Local Similarity 39.6%; Pred. No. 5e-44;
Matches 97; Conservative 41; Mismatches 97; Indels 10; Gaps 5;

QY 190 HCLACG-----KSL-KTPRVVVGEEASVDSWPQVSIQYDKQHVCGGSGILDPHWLTAHC 244
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 176 HC--CGTRRSKTLGOSLRIVGGTEVEGEWPQASLQWDGSHRCGATLINATLWLSAAHC 233
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 245 FRKHTDVFNKVRAGSKLGFPSLAVAKIIIIIEFNPMYFKDNDIALMKLQFLPTFSGTV 304
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 234 FTYKXNPARWTASFQVTIKPSQMKRGLRIIIVHEKYKHPSHDYDISLAELSSPVYTNV 293
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 305 RPICLPFFDEELTPATPLMIIGWFTKQNGKMSDILLQASVQVIDSTRCNADDAYQGEV 364
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 294 HRVCLPDASYEFQPGDVMEVTFGALK-NDGYSQNHRLRQAQVTLIDATTCNEPQYNDAI 352
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 365 TEKXMCAGIPEGGVDTCCGDSGGPLMYQS--DQHVHVGVISWGYGCGGPGSTPGVYTKVSA 422
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 353 TPRMLCAGSLEGKTDACQDGGSGPLVSSDARDIWIYLAGIVSWGDECAKPNKPGVYTRVTA 412
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 423 YLNWI 427
    |||
Db 413 LRDWI 417
    |||

RESULT 12
US-11-265-762-146
/ Sequence 146, Application US/11265762
/ GENERAL INFORMATION:
/ APPLICANT: McCarthy, Sean A.
/ TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
/ FILE REFERENCE: 7853-227-999
/ CURRENT APPLICATION NUMBER: US/11/265,762
/ CURRENT FILING DATE: 2005-11-02
/ PRIOR APPLICATION NUMBER: US/09/796,753
/ PRIOR FILING DATE: 2001-03-01
/ PRIOR APPLICATION NUMBER: 09/183,175
/ PRIOR FILING DATE: 1998-10-30
/ PRIOR APPLICATION NUMBER: 09/223,094
/ PRIOR FILING DATE: 1998-12-30
/ PRIOR APPLICATION NUMBER: 09/223,546
/ PRIOR FILING DATE: 1998-12-30
/ PRIOR APPLICATION NUMBER: 09/224,246
/ PRIOR FILING DATE: 1998-12-30
/ PRIOR APPLICATION NUMBER: 09/259,388
/ PRIOR FILING DATE: 1999-02-26
/ PRIOR APPLICATION NUMBER: 60/122,458
/ PRIOR FILING DATE: 1999-03-01
/ PRIOR APPLICATION NUMBER: 09/312,359
/ PRIOR FILING DATE: 1999-05-14
/ PRIOR APPLICATION NUMBER: 09/336,536
/ PRIOR FILING DATE: 1999-06-18
/ PRIOR APPLICATION NUMBER: 09/342,687
/ PRIOR FILING DATE: 1999-06-29
/ Remaining Prior Application data removed - See File Wrapper or PALM.
```

```
/ NUMBER OF SEQ ID NOS: 162
/ SEQ ID NO 146
/ LENGTH: 423
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-11-265-762-146

Query Match      20.3%; Score 475; DB 7; Length 423;
Best Local Similarity 39.6%; Pred. No. 5e-44;
Matches 97; Conservative 41; Mismatches 97; Indels 10; Gaps 5;

QY 190 HCLACG-----KSL-KTPRVVVGEEASVDSWPQVSIQYDKQHVCGGSGILDPHWLTAHC 244
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 176 HC--CGTRRSKTLGOSLRIVGGTEVEGEWPQASLQWDGSHRCGATLINATLWLSAAHC 233
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 245 FRKHTDVFNKVRAGSKLGFPSLAVAKIIIIIEFNPMYFKDNDIALMKLQFLPTFSGTV 304
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 234 FTYKXNPARWTASFQVTIKPSQMKRGLRIIIVHEKYKHPSHDYDISLAELSSPVYTNV 293
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 305 RPICLPFFDEELTPATPLMIIGWFTKQNGKMSDILLQASVQVIDSTRCNADDAYQGEV 364
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 294 HRVCLPDASYEFQPGDVMEVTFGALK-NDGYSQNHRLRQAQVTLIDATTCNEPQYNDAI 352
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 365 TEKXMCAGIPEGGVDTCCGDSGGPLMYQS--DQHVHVGVISWGYGCGGPGSTPGVYTKVSA 422
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 353 TPRMLCAGSLEGKTDACQDGGSGPLVSSDARDIWIYLAGIVSWGDECAKPNKPGVYTRVTA 412
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 423 YLNWI 427
    |||
Db 413 LRDWI 417
    |||

RESULT 13
US-11-264-096-1294
/ Sequence 1294, Application US/11264096
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: Albumin Fusion Proteins
/ FILE REFERENCE: PF546D1
/ CURRENT APPLICATION NUMBER: US/11/264,096
/ CURRENT FILING DATE: 2005-11-02
/ PRIOR APPLICATION NUMBER: 09/833,245
/ PRIOR FILING DATE: 2001-04-12
/ PRIOR APPLICATION NUMBER: 60/229,358
/ PRIOR FILING DATE: 2000-04-12
/ PRIOR APPLICATION NUMBER: 60/256,931
/ PRIOR FILING DATE: 2000-12-21
/ PRIOR APPLICATION NUMBER: 60/199,384
/ PRIOR FILING DATE: 2000-04-25
/ NUMBER OF SEQ ID NOS: 2267
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 1294
/ LENGTH: 290
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-11-264-096-1294

Query Match      20.0%; Score 467.5; DB 7; Length 290;
Best Local Similarity 37.1%; Pred. No. 1.8e-43;
Matches 99; Conservative 49; Mismatches 92; Indels 27; Gaps 9;

QY 184 GSLVSLHCLACGSKLKTPTRVVVGEEASVDSWPQVSIQYDKQHVCGGSGILDPHWLTAH 243
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16 GSQRAKAATACGRPRMLNRMVGGQDTQEGEWPQVSIQRNGSHFCGGSILIAEQWLTAH 75
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 244 CERKHTDVFNKVRAGSKL-GSFPSSLAVAKIIIEFNPMY---PKDNDIALMKLQFLPT 299
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 76 CFRNTSETSLYQVLLGARQLVQPGPHAMRYARVRQESNPLYQGTASSADVALVEAPVP 135
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 300 FSGTVPICLPFFDEELTPATPL--WIIGWFTKQNGKMSD-----ILLQASVQVIDS 351
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 136 FTNYILPCLP--DPSVIFETGMNCWVTGWSGSPSE-----EDLLPEPILQKLAIPIDT 188
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Qy 352 TRCN-----ADDAQQE-VTEKMMCAGIPGGVDTCCGDSGGPLMYQSDQ-WHVVGIVS 403
Db 189 PKCNLLXSKDTEFGYQPKTIKNDMLCAGFEFGKDKACKGDSGGPLVCLVGLQAGWIS 248
Qy 404 WGYCGGSPSTPGVYTKVSAYLNWYV 430
Db 249 WGECCARQNRGVYIRVTAHHNWIHRI 275

RESULT 14

US-11-268-554-284
; Sequence 284, Application US/11268554
; GENERAL INFORMATION:
; APPLICANT: DOMON, Bruno et al.
; TITLE OF INVENTION: PANCREATIC CANCER TARGETS AND USES
; FILE REFERENCE: THEREOF
; FILE REFERENCE: CL001557
; CURRENT APPLICATION NUMBER: US/11/268,554
; CURRENT FILING DATE: 2005-11-08
; NUMBER OF SEQ ID NOS: 1004
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 284
; LENGTH: 855
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-268-554-284

Query Match 19.8%; Score 463; DB 7; Length 855;
Best Local Similarity 31.5%; Pred. No. 3.1e-42;
Matches 123; Conservative 63; Mismatches 151; Indels 54; Gaps 18;

Qy 58 YFFLCG-----QPLHFIPRKQICDGLDCPLGDEEHCVKSPPEGPANAVRLSKDRSTLQV 113
Db 492 HQFTCKNFKCKPLFWV-----CDSVNDGDNSEQGC-----SCPAQTFRCNGKCC---L 538
Qy 114 LDSATGNWFSACFDNFTALAEACROMGY--SSKPTFAVEIGPDQDLDVVEITENSQE 171
Db 539 SKSQOQNGKDCGDSDEA---SCPKNVVTCTKHTYRC-----LNGLCLSKGNPE 586
Qy 172 LRMNSSGFCLSGLSVLSLHCLACG--KSLKTPRVVVGEEASVDGWPQVSIQ-YDKQHV 228
Db 587 C---DGKEDCSGDS-DEKDC-DCGLRSFTRQARVVGVDADGEWFWQVSLHALGQGHIC 641
Qy 229 GGSILDPHWLTAACHF-----RKHTDVFNMKVVRAG--SDKLGSFPPLSAVAKIIIEFNP 281
Db 642 GASLISPNWLSAAHCYIDDRFRYSPTQWTFGLHDQSQRSAPGVQERRKRLIISHP 701
Qy 282 MPK---DNDIALMKLQPLTFSGTVRPICLPFFDEELTPATPLWIIIGWFTKQNGGKMS 338
Db 702 FFNDFTFDYDIALLEKLEKPAEYSSMVRPICLPDASHVFPAGKAIWVTGWGHT-QYGGTGA 760
Qy 339 DILLQASVQVIDTRCNADDAVQGEVTEKMMCAGIPGGVDTCCGDSGGPLMYQSDQWHV 398
Db 761 LILQGEIRVINTTC---ENLLPQOITPRMCMVGLPSGGVDSQCQDSGGPLSSVEADGRI 818
Qy 399 --VGIVSMGYCGGSPSTPGVYTKVSAYLNWI 427
Db 819 FQAGVWSGDCGAQRNKGVVYTRLPFLFRDWI 849

RESULT 15

US-11-108-459-6
; Sequence 6, Application US/11108459
; GENERAL INFORMATION:
; APPLICANT: Zaas, Aimee
; APPLICANT: Schwartz, David A.
; APPLICANT: Peltz, Gary
; TITLE OF INVENTION: Polymorphic Plasmidogen Genes and Uses Thereof
; FILE REFERENCE: DUKE-09785
; CURRENT APPLICATION NUMBER: US/11/108,459
; CURRENT FILING DATE: 2005-04-18
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent in version 3.3

; SEQ ID NO 6
; LENGTH: 812
; TYPE: PRT
; ORGANISM: Mus musculus
US-11-108-459-6

Query Match 19.6%; Score 459.5; DB 7; Length 812;
Best Local Similarity 39.8%; Pred. No. 7e-42;
Matches 103; Conservative 43; Mismatches 88; Indels 25; Gaps 8;
Qy 181 CLSGSLVSLHCLACGKSLKTP-----RVVGEEASVDGWPQVSI--QYDQHVCGGSIL 233
Db 560 CASAS-----SPEGCKPQVEPKPCGRVVGCVANPHSWPQIULRTFTQGHFCGGTLLI 614
Qy 234 DPHWVLTAAHCFRKHHTDVFNMKVVRAGSDK---LGS-FPSLAVAKIIIEFNPMPYKNDI 289
Db 615 APEWVLTAAHCLKSSRPEFYKVLGAHEEYIRGSDVQEISVAKLI-----LEPNNRDI 668
Qy 290 ALMKLQPLTFSGTVRPICLPFFDEELTPATPLWIIIGWFTKQNGKMSDILLQASVOVI 349
Db 669 ALLKLSRPATITDKVIPAFLSPNMYADRTICVTGNETQGTFG--AGRLKSAQLPVI 726
Qy 350 DSTRCNADDAVQGEVTEKMMCAGIPGGVDTCCGDSGGPLM-YOSDQWHVVGIVSMGYGC 408
Db 727 ENKVCNRVEYLNRRVKSTELCAGQLAGVDSQCQDSGGPLVCFPEKDKYILQGVTSWGLGC 786
Qy 409 GGPSTPGVYTKVSAYLNWI 427
Db 787 ARPKNKPGVYVRVSRFVDWI 805

Search completed: December 5, 2005, 14:04:42
Job time : 14 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2005, 13:42:04 ; Search time 231 Seconds
(without alignments)
1328.592 Million cell updates/sec

Title: US-10-803-530-2
Perfect score: 2342
Sequence: 1 MDPDSQPLSLDKVLRKP.....VYTKSVAYLNIYNNWKAEL 435

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2337	99.8	437	1	TMPS4_HUMAN
2	1844	78.7	435	1	TMPS4_MOUSE
3	762	32.5	388	2	Q4RR7_TETNG
4	705.5	30.1	543	1	TMPSD_MOUSE
5	699.5	29.9	359	2	Q4SPG0_TETNG
6	692.5	29.6	581	1	TMPSD_HUMAN
7	691.5	29.5	486	2	Q5PRA6_BRARE
8	688	29.4	453	2	Q812A6_MOUSE
9	687	29.3	453	1	TMPS3_MOUSE
10	686.5	29.3	767	2	Q9DGR2_XENLA
11	680.5	29.1	722	2	Q6NUF5_XENLA
12	676.5	28.9	492	2	Q6GTK7_HUMAN
13	673.5	28.8	492	1	TMPS2_HUMAN
14	670.5	28.6	492	2	Q96T73_HUMAN
15	660.5	28.2	454	1	TMPS3_HUMAN
16	660.5	28.2	538	2	Q5USC7_HUMAN
17	658	28.1	490	2	Q6P7D7_RAT
18	648.5	27.7	490	1	TMPS2_MOUSE
19	648	27.7	490	2	Q7TN04_MOUSE
20	646	27.6	490	2	Q920K3_RAT
21	636.5	27.2	445	2	Q8CJ17_RAT
22	634	27.1	439	2	Q5RDX7_PONPY
23	631.5	27.0	455	2	Q8CDR0_MOUSE
24	628.5	26.8	455	1	TMPS5_MOUSE
25	612	26.1	371	2	Q8CJ16_RAT
26	585.5	25.0	417	2	Q5R5E8_PONPY
27	584.5	25.0	417	1	HEPS_HUMAN
28	582.5	24.9	1019	1	ENTK_BOVIN
29	578	24.7	436	1	HEPS_MOUSE
30	574	24.5	457	1	TMPS5_HUMAN
31	571	24.4	1035	1	ENTK_BOVIN

32	568.5	24.3	416	1	HEPS_RAT
33	568	24.3	418	1	TM11D_HUMAN
34	552.5	23.6	1069	1	ENTK_MOUSE
35	552	23.6	1034	1	ENTK_PIG
36	550.5	23.5	799	1	TMPS6_MOUSE
37	550.5	23.5	799	2	Q6PF94_MOUSE
38	545	23.3	777	2	Q8CAN9_MOUSE
39	536	22.9	417	1	TM11D_RAT
40	531	22.7	802	1	TMPS6_HUMAN
41	523	22.3	417	1	TM11D_MOUSE
42	518.5	22.1	418	2	Q6IE15_RAT
43	514.5	22.0	730	2	Q4RHT0_TETNG
44	511	21.8	638	1	KLKB1_RAT
45	511	21.8	638	2	Q5FVS2_RAT

ALIGNMENTS

RESULT 1
TMPS4_HUMAN STANDARD; PRT; 437 AA.
ID Q9NRS4: Q5XK06: Q6UX37: Q9NZAS;
AC 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Transmembrane protease, serine 4 (EC 3.4.21.-) (Membrane-type serine protease 2) (MT-SP2).
GN Name=TMPS4; Synonyms=TMPS3; ORFNames=UNQ776/PRO1570;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606; ..
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1).
RC TISSUE=Pancreatic carcinoma;
RX MEDLINE=20283276; PubMed=10825129;
RA Wallrapp C., Haehnel S., Mueller-Pillasch F., Burghardt B., Iwamura T., Ruthenburger M., Lerch M.M., Adler G., Gress T.M.;
RT "A novel transmembrane serine protease (TMPS3) overexpressed in pancreatic cancer.";
RT Cancer Res. 60:2602-2606 (2000).
RN [2]
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1).
SM Smeekens S.S., Lorimer D.D., Wang E., Hou J., Linnevers C.;
RT "MT-SP2, a novel type II membrane serine protease expressed in trachea, colon, and small intestine: identification, cloning, and chromosomal localization.";
RT Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).
TI TISSUE=Colon;
RC PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R., Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H., Sekine M., Ohyashi M., Nishi T., Shibahara T., Tanaka T., Ishii S., Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahara K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K., Tanai H., Kimata M., Watanabe M., Hirakawa S., Chiba Y., Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O., Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S., Yoshioka S., Matsunawa H., Ichihara T., Shiohata N., Sano S., Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O., Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu P., Wakebe H., Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B., Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,

RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hiro M., Ohmori Y.,
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
 RA Matsuura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
 RA Togashi T., Oyama M., Hata H., Shirai Y., Takahashi T., Komatsu T.,
 RA Mizushima-Sugano J., Satoh T., Watanabe M., Komatsu T.,
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.,
 RT "Complete sequencing and characterization of 21,243 full-length human
 RT cDNAs.";
 RL Nac. Genet. 36:40-45 (2004).
 RN [4]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 2).
 RX MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;
 RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D.T., Brush J.,
 RA Chen J., Chow B., Chul C., Crowley C., Currell B., Deuel B., Dowd P.,
 RA Eaton D., Foster J.S., Grimaldi C., Gu Q., Hass P.E., Heldens S.,
 RA Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,
 RA Lewis L., Liao D., Mark M.R., Robbie E., Sanchez C., Schoenfeld J.,
 RA Seshagiri S., Simmons L., Singh J., Smith Y., Stinson J., Vagts A.,
 RA Vandlen R.L., Watanabe C., Wietand D., Woods K., Xie M.-H.,
 RA Yancura D.G., Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A.D.,
 RA Wood W.I., Godowski P.J., Gray A.M.;
 RT "The secreted protein discovery initiative (SPDI), a large-scale
 RT effort to identify novel human secreted and transmembrane proteins: a
 RT bioinformatics assessment.";
 RL Genome Res. 13:2265-2270 (2003).
 RN [5]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORMS 1 AND 3).
 RC TISSUE=Ovary, and Pancreas;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Hellon E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Guimond J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 CC -!- FUNCTION: Probable protease. Seems to be capable of activating
 CC ENAC (By similarity).
 CC -!- SUBCELLULAR LOCATION: Type II membrane protein (Potential).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=3;
 CC Name=1;
 CC IsoId=Q9NRS4-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q9NRS4-2; Sequence=VSP_013117;
 CC Note=No experimental confirmation available;
 CC Name=3;
 CC IsoId=Q9NRS4-3; Sequence=VSP_013116;
 CC Note=No experimental confirmation available;
 CC TISSUE SPECIFICITY: High levels in pancreatic, gastric, colorectal
 CC and ampullary cancer. Very weak expression in normal
 CC gastrointestinal and urogenital tract.
 CC -!- SIMILARITY: Belongs to the peptidase S1 family.
 CC -!- SIMILARITY: Contains 1 LDL-receptor class A domain.
 CC -!- SIMILARITY: Contains 1 peptidase S1 domain.
 CC -!- SIMILARITY: Contains 1 SRCR domain.
 CC -!- CAUTION: Ref.2 sequence differs from that shown due to a

CC CC
 CC frameshift in position 31.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL; AF179224; AAF74526.1; -; mRNA.
 CC EMBL; AF216312; AAF31436.1; -; ALT FRAME; mRNA.
 CC EMBL; AK172766; BADI8749.1; -; mRNA.
 CC EMBL; AY358530; AAQ88894.1; -; mRNA.
 CC EMBL; BC004855; AAQ84855.1; -; mRNA.
 CC EMBL; BC011703; AAH11703.1; -; mRNA.
 CC HSSP; P00760; IEZX.
 CC MEROPS; S01.034; -.
 CC Ensembl; ENSG00000137648; Homo sapiens.
 CC HGNC; HGNC:11878; TMPRSS4.
 CC H-InvDB; HIX0010173; -.
 CC MIM; 606565; -.
 CC GO; GO:0016021; C:integral to membrane; NAS.
 CC GO; GO:0004252; F:serine-type endopeptidase activity; NAS.
 CC GO; GO:0006508; P:proteolysis and peptidolysis; NAS.
 CC InterPro; IPR002172; LDL receptor A.
 CC InterPro; IPR001254; Peptidase S1-S6.
 CC InterPro; IPR001314; Peptidase S1A.
 CC InterPro; IPR001190; Srcr_receptor.
 CC Pfam; PF00089; Trypsin; 1.
 CC PRINTS; PR00722; CHYMOTRYPSIN.
 CC SMART; SM00192; LDLa; 1.
 CC SMART; SM00202; SR; 1.
 CC SMART; SM00202; Tryp_SPC; 1.
 CC PROSITE; PS01209; LDLRA_1; FALSE NEG.
 CC PROSITE; PS00568; LDLRA_2; FALSE NEG.
 CC PROSITE; PS00420; SRCR_1; FALSE NEG.
 CC PROSITE; PS0287; SRCR_2; 1.
 CC PROSITE; PS0240; TRYPSIN_DOM; 1.
 CC PROSITE; PS00134; TRYPSIN_HIS; 1.
 CC PROSITE; PS00135; TRYPSIN_SER; 1.
 CC KW Alternative splicing; Glycoprotein; Hydrolase; Protease;
 CC Serine protease; Signal-anchor; Transmembrane.
 CC TOPO_DOM 1 32 Cytoplasmic (Potential).
 CC TRANSMEM 33 53 Signal-anchor for type II membrane
 CC protein (Potential).
 CC TOPO_DOM 54 437 Extracellular (Potential).
 CC DOMAIN 61 93 LDL-receptor class A.
 CC DOMAIN 94 204 SRCR.
 CC DOMAIN 205 434 Peptidase S1.
 CC ACT_SITE 245 245 Charge relay system (By similarity).
 CC ACT_SITE 290 290 Charge relay system (By similarity).
 CC ACT_SITE 387 387 Charge relay system (By similarity).
 CC SITE 204 205 Cleavage (Potential).
 CC CARBOHYD 130 130 N-linked (GlcNAc...) (Potential).
 CC CARBOHYD 178 178 N-linked (GlcNAc...) (Potential).
 CC DISULFID 64 83 By similarity.
 CC DISULFID 77 92 By similarity.
 CC DISULFID 127 183 By similarity.
 CC DISULFID 140 193 By similarity.
 CC DISULFID 196 310 By similarity.
 CC DISULFID 230 246 By similarity.
 CC DISULFID 356 372 By similarity.
 CC DISULFID 383 410 By similarity.
 CC VARSPLIC 1 3 MLQ -> M (in isoform 3).
 CC FT VARSPLIC 147 151 /FTID=VSP_013116.
 CC FT MISSING (in isoform 2).
 CC FT /FTID=VSP_013117;
 CC FT L -> V (in Ref. 2).
 CC CONFLICT 2 2
 CC SEQUENCE 437 AA; 48204 MW; 351B2FD4A8657B12 CRC64;
 SQ

Query Match 99.8%; Score 2337; DB 1; Length 437;
 Best Local Similarity 100.0%; Pred. No. 3.3e-190;
 Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DPDSQPLNSLDVPLKPRIPMETFRKVGIPPIIIALLSLASIIIVVLLIKVILDKYFL 61
DB 4 DPDSQPLNSLDVPLKPRIPMETFRKVGIPPIIIALLSLASIIIVVLLIKVILDKYFL 63
QY 62 CGQPLHFTPRKQCDGELDCPLGDEDEHCVKSPFEGPAPAVARLSKDRSTLQVLD SATGNW 121
DB 64 CGQPLHFTPRKQCDGELDCPLGDEDEHCVKSPFEGPAPAVARLSKDRSTLQVLD SATGNW 123
QY 122 FSACFDNFTEALAEACRGMYSKPTPRAVEIGPDQDLVDVEITENSQELMRNNSGPC 181
DB 124 FSACFDNFTEALAEACRGMYSKPTPRAVEIGPDQDLVDVEITENSQELMRNNSGPC 183
QY 182 LSGSLVSLHCLACGKSLKTPRVGGEASVDSWPMQVSIQVQKHQVCGSITLDPHWVLT 241
DB 184 LSGSLVSLHCLACGKSLKTPRVGGEASVDSWPMQVSIQVQKHQVCGSITLDPHWVLT 243
QY 242 AHCFKRDHDFVNWKRAGSKLGSPPSLAVAKIIIEFNPMYPKNDIALMKLOPPLTFS 301
DB 244 AHCFKRDHDFVNWKRAGSKLGSPPSLAVAKIIIEFNPMYPKNDIALMKLOPPLTFS 303
QY 302 GTVRPILCLPFDEELTPATPLMIIGWFTKQNGKRWSDILLQASVQVIDSTRCNADDAYQ 361
DB 304 GTVRPILCLPFDEELTPATPLMIIGWFTKQNGKRWSDILLQASVQVIDSTRCNADDAYQ 363
QY 362 GEVTEKMKCAGIPGEGVDTCCGDSGGLMYQSDQHVHVGVISWVGCGGSPTPGVYTKVS 421
DB 364 GEVTEKMKCAGIPGEGVDTCCGDSGGLMYQSDQHVHVGVISWVGCGGSPTPGVYTKVS 423
QY 422 AYLNWYINWVKAEL 435
DB 424 AYLNWYINWVKAEL 437

RESULT 2
ID TMS4 MOUSE STANDARD; PRT: 435 AA.
AC Q8VCA5;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Transmembrane protease, serine 4 (EC 3.4.21.-) (Channel-activating
protease 2) (mCAP2).
DE Name=tmprss4; Synonyms=Cap2;
GN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22144321; PubMed=12149280;
RA Vuagniaux G., Vallet V., Jaeger N.F., Hummler E., Rossier B.C.;
RT "Synergistic activation of ENaC by three membrane-bound channel-
activating serine proteases (mCAP1, mCAP2, and mCAP3) and serum- and
glucocorticoid-regulated kinase (Sgk1) in Xenopus oocytes.";
RL J. Gen. Physiol. 120:191-201(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Mammary tumor;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Probable protease. Seems to be capable of activating
ENaC.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein (Potential).
CC -!- SIMILARITY: Belongs to the peptidase S1 family.
CC -!- SIMILARITY: Contains 1 LDL-receptor class A domain.
CC -!- SIMILARITY: Contains 1 peptidase S1 domain.
CC -!- SIMILARITY: Contains 1 SRCR domain.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL: AY043240; AAK85307.1; -; mRNA.
CC EMBL: BC021368; AAH21368.1; -; mRNA.
CC HSSP: P00760; 1EZX.
CC
CC MEROPS: S01.034; -;
CC Ensembl: ENSMUSG00000032091; Mus musculus.
CC MGI: MGI:2384877; Tmprss4.
CC InterPro: IPR002172; LDL receptor A.
CC InterPro: IPR001254; Peptidase S1 S6.
CC InterPro: IPR001314; Peptidase S1A.
CC InterPro: IPR001190; S1cr_receptor.
CC Pfam: PF00089; Trypsin; 1.
CC PRINTS: PR00722; CHYMOTRYPSIN.
CC SMART: SM00192; LDLA; 1.
CC SMART: SM00202; SR; 1.
CC SMART: SM00202; Tryp_Src; 1.
CC PROSITE: PS01209; LDLRA_1; FALSE_NEG.
CC PROSITE: PS00068; LDLRA_2; FALSE_NEG.
CC PROSITE: PS00420; SRCR_1; FALSE_NEG.
CC PROSITE: PS02087; SRCR_2; 1.
CC PROSITE: PS02040; TRYPSIN_DOM; 1.
CC PROSITE: PS00134; TRYPSIN_HIS; 1.
CC PROSITE: PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Protease; Serine protease; Signal-anchor;
KW Transmembrane.
FT TOPO_DOM 1 30 Cytoplasmic (Potential).
FT TRANSMEM 31 51 Signal-anchor for type II membrane
FT
FT TOPO_DOM 52 435 Extracellular (Potential).
FT DOMAIN 102 202 LDL-receptor class A.
FT DOMAIN 203 432 SRCR.
FT DOMAIN 243 432 Peptidase S1.
FT ACT_SITE 243 243 Charge relay system (By similarity).
FT ACT_SITE 288 288 Charge relay system (By similarity).
FT ACT_SITE 385 385 Charge relay system (By similarity).
FT SITE 202 203 Cleavage (Potential).
FT CARBOHYD 128 128 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 176 176 N-linked (GlcNAc...) (Potential).
FT DISULFID 62 81 By similarity.
FT DISULFID 75 90 By similarity.
FT DISULFID 125 181 By similarity.
FT DISULFID 138 191 By similarity.
FT DISULFID 194 308 By similarity.
FT DISULFID 228 244 By similarity.
FT DISULFID 354 370 By similarity.
FT DISULFID 381 408 By similarity.
SQ SEQUENCE 435 AA; 47496 MW; DC52E45A43E01369 CRC64;
Query Match 78.7%; Score 1844; DB 1; Length 435;
Best Local Similarity 77.2%; Pred No. 3 2e-148;
Matches 336; Conservative 43; Mismatches 56; Indels 0; Gaps 0;
QY 1 MDPSDQPLNSLDVPLKPRIPMETFRKVGIPPIIIALLSLASIIIVVLLIKVILDKYFL 60

Db 1 MESDSGQLNNRDI VPKRRPQETPKKVGIPPIIAVLLSLIALVIVALLIKVILDKYYP 60
Qy 61 LCGOPLHPIPRKQLCDGELDCPLGEDBEHCVKSPFPEGPAVAVRLSKDRSTFLOVLDSATGN 120
Db 61 ICGSPLTFIORGQLCDGHLDCASGEDBEHCVKDFPEKPGVAVRLSKDRSTFLOVLDATGT 120
Qy 121 WFSACFNFTEALAEACRQWYSSKPTFRVIGIGDDODLDVVEITENSQELMRNSSGP 180
Db 121 WASVCFNFTEALAKTACRQWYSDSPAFRAVIRPDNLPVAQVTGNSQELQVQNGRS 180
Qy 181 CLGSSLVSLHCLAGCKSLKTPRVVGGEEASVDSPWQVSIQYDKQHVCGSSILDDEHWLT 240
Db 181 CLGSSLVSLRCLDCGKSLKTPRVVGGVEAPVDSWPWQVSIQYKQHVCGSSILDEHWILT 240
Qy 241 AACFRKHTDVFVWVKRAGSDKLGSPSPSLAVAKIIIEFNPMYPKNDIALMKLOPFTF 300
Db 241 AACFRKLYDVSWKVRAGSNILGNSPLPAKIFIAEPNPLYPKEDIALVKLQWLT 300
Qy 301 SGTVRPCLPFFDEELTPATPLWIGWFTKONGKMSDILLQASVOVLDSTRCNADDAY 360
Db 301 SGTVRPCLPFFDEELTPATPLWIGWFTKONGKMSDILLQASVOVLDSTRCNADDAY 360
Qy 361 QGVTETKMCAGIPEGVDTCQDGGPLMYQSDQVHVGVISWGGCGPSTPGVYTKV 420
Db 361 EGEVTAEMLCAGTPGGKDTCCQDGGPLMYHSDKWQVGVISWGGCGPSTPGVYTKV 420
Qy 421 SAYLNNIYNVWKAEL 435
Db 421 TAYLNNIYNVRKSEM 435

RESULT 3
Q4RRR7 TETNG
ID Q4RRR7 TETNG PRELIMINARY; PRT; 388 AA.
AC Q4RRR7;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome 16 SCAP15002, whole genome shotgun sequence.
DE (Fragment)
GN ORFNames=GSTENG00030043001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jallion O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bounneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Bionnet C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.,
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
the early vertebrate proto-karyotype.",
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; CAAE01015002; CAG08915.1; -; Genomic_DNA.
CC

FT NON TER 1
FT NON_TER 388
SQ SEQUENCE 388 AA; 42696 MW; 07B0ADF4BA4C5736 CRC64;
Query Match 32.5%; Score 762; DB 2; Length 388;
Best Local Similarity 39.3%; Pred. No. 3.8e-56;
Matches 151; Conservative 56; Mismatches 121; Indels 56; Gaps 8;
Qy 95 PEGPAVAVRLSKDRSTFLOVLDATGNSWFSACFNFTEALAEACRQWYSSKPTFRV-- 152
Db 13 PHSGSVFYRISPESSLEIQGLKPLTWLPVCYERWNSSLGTLVCRQLGYLRLTKHGVL 72
Qy 153 -EIGDDODLDVEIT-----ENSQELMRNSSGPGCLSGSLVSLHCLAGCKSLKTPRV 204
Db 73 TDIGPNYTDGFIQITPEHNSLENIWQLR-----GSCVTGKVLSQLCFECGTRAKLPRII 127
Qy 205 GGEASVDSPWQVSIQYDKQHVCGSSILDPHWLTAAHCFRKHTD----- 250
Db 128 GGVEATLGRWPQVSLYSSRHTCGGSIINSQWVTAACHVAVSSLVGLSPNRSLS 187
Qy 251 -----VFNWVRA-----GSDKLGSPSPSLAVAKIIIEFNPMY---PKDN 287
Db 188 TSVLSEMSATGVLRSVSWVYAGIITRGSAKVAEHVGYAVEKII---YNKEYNHRSHDG 244
Qy 288 DIALMKLQFELTSGTVRPICLPFFDEELTPATPLWIGWFTKONGKMSDILLQASVQ 347
Db 245 DIALKLRTLFNFDIIRPVCLPQYDYEPPGGTQCWISGWYTPQEGVHSPDTLKEAPVP 304
Qy 348 VIDSTRCNADDAYQGEVTEKMCAGIPEGVDTCQDGGPLMYQSDQ-VHVGVISWGY 406
Db 305 IISTKRCNSCMYNGEITSRMLCAGYTEKGKACQDGGPLVCCQDENVWRLAGVVSWS 364
Qy 407 GCGGPGTPGVYTKVSAVLNNIYNV 430
Db 365 GCAEPNHPGVYTKVAEFLGWIYDM 388

RESULT 4
TMPSD_MOUSE
ID TMPSD_MOUSE STANDARD; PRT; 543 AA.
AC QSU405; Q8CFE0; Q91VQ8;
DT 10-MAY-2005 (Rel. 47, Created)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Transmembrane protease, serine 13 (EC 3.4.21.-) (Mosaic serine
protease) (Membrane-type mosaic serine protease).
GN Name=Tmpsr13; Synonyms=Msp;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC STRAIN=B5/EGFP, and FVB/N;
RC TISSUE=Mammary tumor, and Trophoblast stem cell;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan B., Moore T., Max S.I., Wang J., Hong L.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Heiton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,

DR EMBL; AB048977; BAB39742.1; -; mRNA.
 DR EMBL; AY190317; AAB38062.1; -; mRNA.
 DR EMBL; AK027798; BAB55376.1; -; mRNA.
 DR EMBL; AP002962; -; NOT_ANNOTATED_CDS; Genomic_DNA.
 DR HSSP; P00760; 1E2X.
 DR MEROPS; S01.087; -.
 DR Ensembl; ENSG00000137747; Homo sapiens.
 DR HGNC; HGNC:29808; TMRS313.
 DR GO; GO:0016021; C:integral to membrane; NAS.
 DR GO; GO:0004252; F:serine-type endopeptidase activity; NAS.
 DR GO; GO:0006508; F:proteolysis and peptidolysis; NAS.
 DR InterPro; IPR002172; LDL receptor A.
 DR InterPro; IPR001254; Peptidase S1_S6.
 DR InterPro; IPR001134; Peptidase_S1A.
 DR InterPro; IPR001190; Srcr_receptor.
 DR Pfam; PF00089; Trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00020; Tryp SPG; 1.
 DR PROSITE; PS01209; LDLRA_1; FALSE_NEG.
 DR PROSITE; PS00668; LDLRA_2; FALSE_NEG.
 DR PROSITE; PS00420; SRCR_1; FALSE_NEG.
 DR PROSITE; PS0287; SRCR_2; 1.
 DR PROSITE; PS0240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Alternative splicing; Glycoprotein; Hydrolase; Polymorphism; Protease;
 KW Repeat; Serine protease; Signal-anchor; Transmembrane.
 FT TOPO_DOM 1 160
 FT TRANSMEM 161 181
 FT TOPO_DOM 182 581
 FT REPEAT 9 13
 FT REPEAT 14 18
 FT REPEAT 19 23
 FT REPEAT 24 28
 FT REPEAT 29 33
 FT REPEAT 34 38
 FT REPEAT 39 43
 FT REPEAT 44 48
 FT REPEAT 49 53
 FT REPEAT 54 58
 FT REPEAT 59 63
 FT REPEAT 64 68
 FT REPEAT 69 73
 FT REPEAT 74 78
 FT REPEAT 79 83
 FT REPEAT 84 88
 FT REPEAT 89 93
 FT DOMAIN 190 320
 FT DOMAIN 199 221
 FT REGION 9 88
 FT REGION 14 68
 FT COMPBIAS 9 89
 FT ACT_SITE 361 361
 FT ACT_SITE 409 409
 FT ACT_SITE 506 506
 FT CARBOHYD 250 250
 FT CARBOHYD 287 287
 FT CARBOHYD 400 400
 FT CARBOHYD 440 440
 FT DISULFID 245 309
 FT DISULFID 258 312
 FT DISULFID 346 362
 FT DISULFID 443 512
 FT DISULFID 475 491
 FT DISULFID 502 530
 FT VARSPLIC 146 181
 FT VARSPLIC 423 486
 FT EQVNLDFKKNNDLVDSYLT -> GEGICTPSRAPQP
 FT QHPLOPSHLSASVNSYTPKASAGQSKTLKDPYMEHFCFI
 FT IRETAQGL (in isoform 4).

FT VARSPLIC 487 581
 FT VARSPLIC 555 581
 FT VARSPLIC 555 558
 FT VARSPLIC 559 581
 FT CONFLICT 187 187
 FT CONFLICT 192 192
 FT CONFLICT 201 201
 FT CONFLICT 206 206
 FT CONFLICT 259 259
 FT CONFLICT 298 298
 FT CONFLICT 496 496
 SQ SEQUENCE 581 AA; 62680 MW; A4935CCCF31D29EE CRC64;
 Query Match 29.6%; Score 692.5; DB 1; Length 581;
 Best Local Similarity 35.2%; Pred. No. 5.5e-50;
 Matches 151; Conservative 70; Mismatches 161; Indels 47; Gaps 11;
 QY 22 IPMETER--KVGIPITIALLSASIIIVVLLIKVLDKYVFLCGQ-----PLH 67
 Db 149 LPKFTWREGQQLPLIGCVLLLIALLVSLIIL-----FQWQGHGTGIRYKEQRESCPH 202
 QY 68 FIPRKLQCDGELDCPLGDEDEHCVKSPFEGPAVAVRLSKDRSTLQVLDSATGNWFSACFD 127
 Db 203 AV----RCDGVVCKLKSDELGC-----VRPDWKSLLKIYSGSHQWLPICSS 247
 QY 128 NFTAALAEACRQMGYSKPTFRAVEIGPDQDLVVEITENSQELMRNMSGPCLSGSLV 187
 Db 248 NWNDSYSEKTCQQLGFES--AHRTEVAHRDFANSFSLRYNSTIQESLHRSECPQRYI 305
 QY 188 SLHCLACGKSLKTPRVVGEASVDSPWQVSIQVOKHVCVGGSTLDPHWLTAACF-- 245
 Db 306 SLQCSHCLRAMTRIGVGGALASDKPWQVSLHFGTHICGGLIDAQWLTAACFFV 365
 QY 246 RKHTDVFNVKVRAGSKLGSFPLA--VAKIIIIIFENPMY---PKNDIALMKLOPLTFPS 301
 Db 366 TREKVLGKWKVYAGTSNLHLPEAAIAEII---NSNYTDEDDYDIALMRLSKPLTSL 422
 QY 302 GTVRPICLPFFDEBLTPATPLWIIWGFTKONGKMSDILLOASVQVVDSTRCNADDAVQ 361
 Db 423 AHIHPACLPMHGQTFSLNETCWTGFGKTRTDDKTSPLREVQVNLIDFKKCNLYVD 482
 QY 362 GEVTEKAMCAGIPSGVDVTCQDSGGPLM--YQSDQWVVGIVSGYGGGPGSTPGVYTKV 420
 Db 483 SYLTPRMKAGDLHGGDRSCQDSGGPLVCQNNRWYLAGVTSWGTGCGQRNKPQVYTKV 542
 QY 421 SAYLNWIYN 429
 Db 543 TEVLFWIYS 551
 RESULT 7
 QSPRA6 BRARE PRELIMINARY; PRT; 486 AA.
 ID Q5PRA6 BRARE PRELIMINARY; PRT; 486 AA.
 AC Q5PRA6; 29, Created
 DT 01-FEB-2005 (Tremblrel. 29, Last sequence update)
 DT 01-FEB-2005 (Tremblrel. 29, Last annotation update)
 DE Zgc:101791.
 GN ORFNames=zgc:101791;
 OS Brachydanio rerio (zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OC NCBI_TaxID=7955;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Olfactory epithelium;


```
Db 143 AOLGPPSYSSDHLRVDALEBQFQDFVSNHLLSDDKVTALHHSVYMRG---CTS GHV 199
Qy 187 VSLHCLACG-KSLKTPRVVGGBEASVDSPWQVSIQYDKQHVCGGSIILDPHWLTAACHF 245
Db 200 VTLKCSACGTRGYSPRIVGNNSSLTQWPWQVSLQFQGYHLCCGSIITPLVITAAHCV 259
Qy 246 RKHTDVFNNKVRAGSKL--GSFPLSLAVAKIIIEFNPMY-PK--DNDIALMKLOPLTF 300
Db 260 YDLHPKSWTVQVGLVSLMDSVPVSHLVEKII---YHSKYKPKRLGNDIALMKLSEPLTF 316
Qy 301 SGTVPRIPLPFFDELTPTALWIIGWFTKQNGKMSDILLQASVQVIDSTRCNADDAY 360
Db 317 DETIQICLPNSEENFPDKLWTSWGAT-EDGDASPLVNAHAAVPLISNKNHRDYY 375
Qy 361 QGEVTEKMKCAGIPGEGVDTCCGSGGGLMYQSDQ-WHVVGIVSWGYGCGGSTPGVYTK 419
Db 376 GGIISPSMLCAGYLKGGVDCQDGGPLVLCQERRLWKLVGATSPGIGCAEVNKPQVYTR 435
Qy 420 VSAEYLNWYVNVKAEI 435
Db 436 ITSFLDWIHEQLERDL 451

RESULT 9
TPRS3 MOUSE STANDARD; PRT; 453 AA.
AC Q8KIT0; Q8VDE0;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE 10-MAY-2005 (Rel. 47, Last annotation update)
DE Transmembrane protease, serine 3 (EC 3.4.21.-).
GN Name=TPRS3;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE, SUBCELLULAR LOCATION, AND FUNCTION IN ENAC
CLEAVAGE.
RX MEDLINE=22281255; PubMed=12393794; DOI=10.1093/bmg/11.23.2829;
RA Guipponi M, Vuagniaux G, Wattenhofer M, Shibuya K, Vazquez M.,
RA Dougherty L., Scamuffa N., Guida E., Okui M., Rossier C., Hancock M.,
RA Buchet K., Reymond A., Hummler E., Marzella P.L., Kudoh J.,
RA Shimizu N., Scott H.S., Antonarakis S.E., Rossier B.C.;
RT "The transmembrane serine protease (TPRS3) mutated in deafness
RT DFNB8/10 activates the epithelial sodium channel (ENAC) in vitro.";
RL Hum. Mol. Genet. 11:2829-2836(2002).
CC -!- FUNCTION: Probable protease. Seems to be capable of activating
CC ENAC.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Endoplasmic
CC reticulum.
CC -!- TISSUE SPECIFICITY: Expressed in the spiral ganglion, the cells
CC supporting the organ of Corti and the stria vascularis.
CC -!- PTM: Undergoes autoproteolytic activation.
CC -!- SIMILARITY: Belongs to the peptidase S1 family.
CC -!- SIMILARITY: Contains 1 LDL-receptor class A domain.
CC -!- SIMILARITY: Contains 1 peptidase S1 domain.
CC -!- SIMILARITY: Contains 1 SRCR domain.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC -----
CC EMBL; AJ242916; CAD22137.1; -; Genomic DNA.
CC DR EMBL; AJ300738; CAC83350.1; -; mRNA.
CC DR HSSP; P00760; 1EZX.
CC DR MEROPS; S01.079; -.
CC
CC Ensembl; ENSMUSG0000024034; Mus musculus.
CC MGI; MGI:2155445; Tprss3.
CC GO; GO:0016021; C:integral to membrane; TAS.
```

```
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR001254; Peptidase_S1_S6.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR001190; Srcr_receptor.
DR Pfam; PF00057; Ldl_recept_a; 1.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PROSITE; PS01209; LDLRA_1; 1.
DR PROSITE; PS00068; LDLRA_2; 1.
DR PROSITE; PS00287; SRCR_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Endoplasmic reticulum; Glycoprotein; Hydrolase; Protease;
KW Serine protease; Signal-anchor; Transmembrane; zymogen.
FT TOPO_DOM 1 48 Cytoplasmic (Potential).
FT TRANSMEM 49 69 Signal-anchor for type II membrane protein (Potential).
FT TOPO_DOM 70 453 Extracellular (Potential).
FT DOMAIN 72 108 LDL-receptor class A.
FT DOMAIN 104 205 SRCR.
FT DOMAIN 217 448 Peptidase S1.
FT ACT_SITE 257 257 Charge relay system (By similarity).
FT ACT_SITE 304 304 Charge relay system (By similarity).
FT ACT_SITE 400 400 Charge relay system (By similarity).
FT SITE 216 217 Cleavage (Potential).
FT CARBOHYD 221 221 N-linked (GlcNAc...) (Potential).
FT DISULFID 73 85 By similarity.
FT DISULFID 79 98 By similarity.
FT DISULFID 92 107 By similarity.
FT DISULFID 129 194 By similarity.
FT DISULFID 142 204 By similarity.
FT DISULFID 207 324 By similarity.
FT DISULFID 242 258 By similarity.
FT DISULFID 338 406 By similarity.
FT DISULFID 369 385 By similarity.
FT DISULFID 396 424 By similarity.
FT CONFLICT 117 117 L -> H (in Ref. 1; CAC83350).
FT CONFLICT 246 246 V -> I (in Ref. 1; CAC83350).
SQ SEQUENCE 453 AA; 49492 MW; 1ABCBF10AF6E1EF6 CRC64;

Query Match 29.3%; Score 687; DB 1; Length 453;
Best Local Similarity 36.9%; Pred. No. 1.1e-49;
Matches 161; Conservative 72; Mismatches 153; Indels 50; Gaps 14;

Qy 22 IPMTFRKVGPIIILALLSLASIIIVVLIKVILD---KYFLCQGPLHFIPRKQLCDGE 78
Db 44 LPLKEFPPIIVIGIILALALA-----ICLGHFDGCKY--RCHSSPKCIELTARCDGV 95
Qy 79 LDCPLGDEEHCVKSPFEGPAVAVRLSKDRSTLQVLDSATGNWFSACDFNFTALAEATC 138
Db 96 SDCKNADEYRC-----VRVSGRAALQVFTAAA--WRTWCSDDMKSHYAKIAC 142
Qy 139 ROMGYSS-----KPTFRAVEIGPDOLDVVEITENSQELWRNSSGCPCLSGSL 186
Db 143 AOLGPPSYSSDHLRVDALEBQFQDFVSNHLLSDDKVTALHHSVYMRG---CTS GHV 199
Qy 187 VSLHCLACG-KSLKTPRVVGGBEASVDSPWQVSIQYDKQHVCGGSIILDPHWLTAACHF 245
Db 200 VTLKCSACGTRGYSPRIVGNNSSLTQWPWQVSLQFQGYHLCCGSIITPLVITAAHCV 259
Qy 246 RKHTDVFNNKVRAGSKL--GSFPLSLAVAKIIIEFNPMY-PK--DNDIALMKLOPLTF 300
Db 260 YDLHPKSWTVQVGLVSLMDSVPVSHLVEKII---YHSKYKPKRLGNDIALMKLSEPLTF 316
Qy 301 SGTVPRIPLPFFDELTPTALWIIGWFTKQNGKMSDILLQASVQVIDSTRCNADDAY 360
Db 317 DETIQICLPNSEENFPDKLWTSWGAT-EDGDASPLVNAHAAVPLISNKNHRDYY 375
Qy 361 QGEVTEKMKCAGIPGEGVDTCCGSGGGLMYQSDQ-WHVVGIVSWGYGCGGSTPGVYTK 419
Db 376 GGIISPSMLCAGYLKGGVDCQDGGPLVLCQERRLWKLVGATSPGIGCAEVNKPQVYTR 435
```

```
QY 420 VSAYLNWIYNVWKAEL 435
Db 436 ITSFLDWIHEQLERDL 451

RESULT 10
Q9DGR2_XENLA
AC Q9DGR2_XENLA PRELIMINARY; PRT; 767 AA.
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Embryonic serine protease-2.
GN Name=Xsp-2;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20363741; PubMed=10903452; DOI=10.1016/S0378-1119(00)00225-0;
RA Yamada K., Takabatake T., Takeshima K.;
RT "Isolation and characterization of three novel serine protease genes
from Xenopus laevis."
RL Gene 252:209-216 (2000).
DR EMBL; AB038497; BAB08217.1; -; mRNA.
DR HSP; P00760; IEZX.
DR MEROPS; S01.049; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0005044; F:scavenger receptor activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR001254; Peptidase_S1A.
DR InterPro; IPR001130; Srcr_receptor.
DR Pfam; PF00057; Ldl_recept_4.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00261; LDLRECEPTOR.
DR SMART; SM00192; LDLA; 8.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS01209; LDLRA_1; 8.
DR PROSITE; PS00668; LDLRA_2; 2.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Protease.
SQ SEQUENCE 767 AA; 96001 MW; E0566A38796DE96E CRC64;

Query Match 29.3%; Score 686.5; DB 2; Length 767;
Best Local Similarity 38.3%; Pred. No. 2.6e-49;
Matches 145; Conservative 58; Mismatches 153; Indels 23; Gaps 7;

QY 62 CQQLHFTIPRKQLCDGELDCLPEDEHCVKSPPEGPAVAVRLSKDRSTLQVLDSATGNW 121
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
395 CGSSVSCVLLSQWCDGVDCPYGDEMCSVLYPAD-----FQLQVYTSVSAAW 443

QY 122 FSACFDNFTEALATACRMQYSKPTFRAVEI-----GPDDLDVVEITENSQELMRNS 177
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
444 LPVCSDFWDDDFGRFACQDFGNGSSYNNRYDTLMSYPAPNGYFKLYSGYWRSKFYTSVQY 503

QY 178 SGPLSGSLSLHCLACGKSLKT--PRVVGEEASVDSPWQVSIQYDKHVCVGGSIILDP 235
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
504 SSYCVSGNVVSLHCISGVNSNLSVSRIVGTTFANLGNWPNQVLYITGVLCGGSIIIF 563

QY 236 HWVLTAAHC-FRKHTVFNWKNVRAGSKLGSFFSLA---VAKIIIEFNPMYPKNDNDIAL 291
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
564 KWIIVTAACHVGYSSYSSASGRVFAGLTKPSYNNASAYFVERIIVHPGYKSTYNDNDIAL 623
```



```
DR InterPro: IPR001254; Peptidase S1_S6.
DR InterPro: IPR001190; Serr_receptor.
DR Pfam: PF00057; Ldl_recept_a; 4.
DR Pfam: PF00089; Trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00261; LDLRECEPTOR.
DR SMART: SM00192; LDLA; 7.
DR SMART: SM00202; SR; 1.
DR SMART: SM00020; Tryp_SPC; 1.
DR PROSITE: PS01209; LDLRA1; 7.
DR PROSITE: PS00068; LDLRA2; 1.
DR PROSITE: PS0287; SRCK2; 1.
DR PROSITE: PS0240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
DR Hypothetical protein.
SQ SEQUENCE 722 AA; 80745 MW; 99D2B7391AE983FC CRC64;

Query Match 29.1%; Score 680.5; DB 2; Length 722;
Best Local Similarity 38.5%; Pred. No. 7.7e-49;
Matches 146; Conservative 57; Mismatches 153; Indels 23; Gaps 7;

QY 62 CQQLHFIPIKQCDGLDCLPGEDEHCHVKSPFPGPAVAVRLSKDRSTLQVLDSATGNW 121
DB 350 CGSSVSCVLSSQWCDVSDCPYGEDMSC-----VRLYGADFQLQVYITSVAW 398
QY 122 FSACDNFTALAEATACRQMGYSKPTFRAVEI---GPDQDLVDVEITENSQELMRNS 177
DB 399 LPVCSYDNDPFRACQDFGYSNGSYNYRDTLMSYPAPNGYFKLYSGWRSKFTYSQY 458
QY 178 SGPLSGSLVSLHLCACGSKLKT--PRVVGGEASVDSMPQVSIQYDKQHVCGGSILDP 235
DB 459 SSYCYSGNVVSLHLCISGVSYSVSRIVGGTFANLGNWPMQVNIQLITGVLCGGSIIIP 518
QY 236 HWLTLAHC-FRKHTDVFNWVKRAGSKLGSPPSLA---VAKIIIEFNPMYPKNDIAL 291
DB 519 KWIIVTAAHCHVYSGSASGRVFAGLTLTKPSYNASAYFVERIIIVHPGYKSYTYNDIAL 578
QY 292 MKLQPLFTSGTVRPICLPFPDEELTPATPLWIGMGFTKQNGKMSDILLQASVOVIDS 351
DB 579 MKLRDEITFGYTTQVPCLPNSGMFWAAGTTWISGWSTYE--GGSVSTYVQLVAAIPLIDS 637
QY 352 TRCNADDAAYQGEVTEKMKCAGIPEGGVDTCCQDSGGPLM-YOSDQHWVVGIVSWGVC GG 410
DB 638 NVNCQSVYVNGOITSSMICAGYLSGVDTCCQDSGGPLVKNKNGTWLVGDTSWGDCAR 697

QY 411 PSTPGYTKVSAYLNIYN 429
DB 698 ANKPGVYGNVTFLEWIYS 716

RESULT 12
Q6GTK7 HUMAN
ID Q6GTK7 HUMAN PRELIMINARY; PRT; 492 AA.
AC Q6GTK7.
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Transmembrane protease, serine 2 variant.
GN Name=TWPRSS2;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Peripheral Nervous System;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
```

```
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Borak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN (2)
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Peripheral Nervous System;
RA Strausberg R.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
RN (3)
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RA Maruyama K., Sugano S.;
RT "Oligo-capping: a simple method to replace the cap structure of
RT eucaryotic mRNAs with oligoribonucleotides.";
RL Gene 138:171-174(1994).
RN (4)
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RA Suzuki Y., Yoshitomo K., Maruyama K., Suyama A., Sugano S.;
RT "Construction and characterization of a full length-enriched and a 5'-
RT end-enriched cDNA library.";
RL Gene 200:149-156(1997).
RN (5)
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RA Suzuki Y., Sugano S., Totoki Y., Toyoda A., Takeda T., Sakaki Y.,
RA Tanaka A., Yokoyama S.;
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC051839; AAHS1839.1; -; mRNA.
DR EMBL; AK222784; BAD96504.1; -; mRNA.
KW Protease; Transmembrane.
SQ SEQUENCE 492 AA; 53859 MW; C05B5531C8A311C7 CRC64;

Query Match 28.9%; Score 676.5; DB 2; Length 492;
Best Local Similarity 39.1%; Pred. No. 1e-48;
Matches 150; Conservative 57; Mismatches 128; Indels 49; Gaps 13;

QY 75 CDGELDCPLGDEDEHCHVKSPFPGPAVAVRLSKDRSTLQVLDSATGNWFSACFDNFEALA 134
DB 133 CDGVSHCFGDEENRCVRLY--GP-----NFILQVYSQRKSWHPVQDDWNYNG 181
QY 135 ETACRQMGYSKPTFRAVEIIGPDQDLVDVEITENSQELMRNSG-----PCL 182
DB 182 RAACRDMGY--KNFYSSQ-----GIVDSGSGTSFMKLTNSAGNVDIYKLYHSDACS 232
QY 183 SGSLVSLHLCACGSKL---KTPRVVGGEASVDSMPQVSIQYDKQHVCGGSILDPHWL 239
DB 233 SKAVVSLRFLACIACGVNLSRSQSRIVGGSALPGAWPQVSLHVQNVHVGCGSIITPEIV 292
QY 240 TAAHCFRKH-TDVFNWKVRAGSKLGSF-----PSLAVALAKIIIEFNPMY---PKNDIAL 291
DB 293 TAAHCVEKPLNNPWHWTAFAGILR-QSFMFYGAGYQVEKVI---SHPNVDSTKNDIAL 348
QY 292 MKLQPLFTSGTVRPICLPFPDEELTPATPLWIGMGFTKQNGKMSDILLQASVOVIDS 351
DB 349 MKLQPLFTFNDLVKVPCLPNSGMFWAAGTTWISGWSTYE--GGSVSTYVQLVAAIPLIDS 407
QY 352 TRCNADDAAYQGEVTEKMKCAGIPEGGVDTCCQDSGGPLM-YOSDQHWVVGIVSWGVC GG 410
DB 408 QRCNRYVDNLITPAMTCAGFLQGVNDSQCGSGPLVTSKNIWLVIGTWSGSGCAK 467
```

```

OY 411 PSTPGVYTKVAYLNWYVNWAE 434
Db 468 AYRPGVIGNVWFTDWMQMRAD 491

RESULT 13
TMPRSS2_HUMAN
ID TMPRSS2_HUMAN STANDARD; PRT; 492 AA.
AC O15393; Q9EXX1;
DT 15-JUL-1998 (Rel. 36, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Transmembrane protease, serine 2 precursor [EC 3.4.21.-] [Contains:
DE Transmembrane protease, serine 2 non-catalytic chain; Transmembrane
DE protease, serine 2 catalytic chain].
GN Name=TMPRSS2; Synonyms=PRSS10;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RN NUCLEOTIDE SEQUENCE.
RP MEDLINE=97468144; PubMed=9325052; DOI=10.1006/geno.1997.4845;
RA Paoloni-Giacobino A., Chen H., Peitsch M.C., Rossier C.,
RA Antonarakis S.E.
RT "Cloning of the TMPRSS2 gene, which encodes a novel serine protease
RT with transmembrane, LDLRA, and SRCR domains and maps to 21q22.3."
RL Genomics 44:309-320(1997).
RN [2]
RN NUCLEOTIDE SEQUENCE.
RP MEDLINE=21309069; PubMed=11414763; DOI=10.1006/geno.2001.6551;
RA Teng D.-H., Chen Y., Lian L., Ha P.C., Tavtigian S.V., Wong A.K.C.;
RT "Mutation analyses of 268 candidate genes in human tumor cell lines."
RL Genomics 74:352-364(2001).
RN [3]
RN NUCLEOTIDE SEQUENCE, AND MUTAGENESIS.
RP MEDLINE=21139112; PubMed=11245484;
RA Afar D.E.H., Vivanco I., Hubert R.S., Kuo J., Chen E., Saffran D.C.,
RA Raitano A.B., Jakobovits A.;
RT "Catalytic cleavage of the androgen-regulated TMPRSS2 protease results
RT in its secretion by prostate and prostate cancer epithelia."
RL Cancer Res. 61:1686-1692(2001).
RN [4]
RP TISSUE SPECIFICITY.
RX MEDLINE=21104370; PubMed=11169526;
RA Vaarala M.H., Porvari K.S., Kellokumpu S., Kyllonen A.P., Vihko P.T.;
RT "Expression of transmembrane serine protease TMPRSS2 in mouse and
RT human tissues."
RL J. Pathol. 193:134-140(2001).
CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Activated by
CC cleavage and secreted.
CC -!- TISSUE SPECIFICITY: Expressed strongly in small intestine. Also
CC expressed in prostate, colon, stomach and salivary gland.
CC -!- SIMILARITY: Belongs to the peptidase S1 family.
CC -!- SIMILARITY: Contains 1 LDL-receptor class A domain.
CC -!- SIMILARITY: Contains 1 peptidase S1 domain.
CC -!- SIMILARITY: Contains 1 SRCR domain.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; U75329; AAC51784.1; -; mRNA.
CC EMBL; AF123453; AAD37117.1; -; mRNA.
CC EMBL; AF270487; AAK29280.1; -; mRNA.
CC HSP; P00760; 1EZ3.
CC MEROPS; S01.247; -.
CC
CC Ensembl; ENSG00000184012; Homo sapiens.
CC HGNC; HGNC:11876; TMPRSS2.

```

```

DR MIM; 602060; -.
DR GO; GO:0005887; C:integral to plasma membrane; TAS.
DR GO; GO:0008236; F:serine-type peptidase activity; TAS.
DR InterPro; IPR001272; LDL receptor A.
DR InterPro; IPR001254; Peptidase_S1_S6.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR001190; Srcr_receptor.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00192; LDLa; 1.
DR SMART; SM00202; SR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS01209; LDLRA_1; 1.
DR PROSITE; PS00668; LDLRA_2; 1.
DR PROSITE; PS00420; SRCR_1; FALSE_NEG.
DR PROSITE; PS02087; SRCR_2; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Polymorphism; Protease; Serine protease;
KW Signal-anchor; Transmembrane; Zymogen.
FT CHAIN 1 255
FT Transmembrane protease, serine 2 non-
FT catalytic chain.
FT CHAIN 254 492
FT Transmembrane protease, serine 2
FT catalytic chain.
FT TOPO_DOM 1 84
FT TRANSMEM 85 105
FT Cytoplasmic (Potential).
FT Signal-anchor for type II membrane
FT protein (Potential).
FT FT 106 492
FT Extracellular (Potential).
FT FT 112 149
FT LDL-receptor class A.
FT FT 150 242
FT SRCR.
FT FT 256 489
FT Peptidase S1.
FT FT 296 296
FT Charge relay system (By similarity).
FT FT 345 345
FT Charge relay system (By similarity).
FT FT 441 441
FT Charge relay system.
FT FT 255 256
FT Cleavage (Potential).
FT FT 213 213
FT N-linked (GlcNAc...) (Potential).
FT FT 249 249
FT N-linked (GlcNAc...) (Potential).
FT FT 113 126
FT By similarity.
FT FT 120 139
FT By similarity.
FT FT 133 148
FT By similarity.
FT FT 172 231
FT By similarity.
FT FT 185 241
FT By similarity.
FT FT 244 365
FT Interchain (between non-catalytic and
FT catalytic chains) (By similarity).
FT FT 281 297
FT By similarity.
FT FT 410 426
FT By similarity.
FT FT 437 465
FT By similarity.
FT FT 449 449
FT K -> N (in dbSNP:1056602).
FT FT 255 255
FT /FTID=VAR 011692.
FT FT 441 441
FT R -> Q: Loss of cleavage.
FT FT 160 160
FT S -> A: Loss of activity.
FT FT 242 242
FT M -> V (in Ref. 3).
FT FT 329 329
FT I -> L (in Ref. 1).
FT FT 489 491
FT E -> Q (in Ref. 1).
FT FT 492 AA; 53891 MW; CAB44FD174A9076B CRC64;
SQ SEQUENCE 492 AA; 53891 MW; CAB44FD174A9076B CRC64;

Query Match 28.8%; Score 673.5; DB 1; Length 492;
Best Local Similarity 38.8%; Pred.No.1.8e-48;
Matches 149; Conservative 58; Mismatches 128; Indels 49; Gaps 13;

OY 75 CDGELDCPLGEDEHCVKSFPEGPAPAVAVRLSKDRSTLQVLDSATGNWFSACFDNFTEALA 134
Db 133 CDGVSHCPGGEDEHCVRLY--GP-----NFIQMYSSQKSWHPVCQDDWNNYVG 181
OY 135 ETACRQMGYSKPTFAVEIGPDQLDVVEITENQELMRNSSG-----PCL 182
Db 182 RAACRDMGY--KXNFYSQ-----GIVDSDSGTSFMKLNTSAGNVDYTKLYHSDACS 232
OY 183 SGLSVSLHCLACKSL---KTPRVVGGEEASVDSWPVOVSIQYDKQHVCGGSIIDPHWVL 239
Db 233 SKAVVSLRCIACGVNLNSRQSRIVGGESALPGANFWQVSLHVQNVHVCVCGSIITPEWIV 292

```

```
Qy 240 TAAHCFRKH-TDVFVWVWVAGSDKLGSP-----PSLAVAKIIIIIEFNPMY---PKNDIAL 291
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 293 TAAHCFRKH-TDVFVWVWVAGSDKLGSP-----PSLAVAKIIIIIEFNPMY---PKNDIAL 348
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 292 MKLQPLTFSGTVRPICLPFDEBELTPATPLWILGWFTKONGKMSDILLOASVOVIDS 351
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 349 MKLQPLTFNDLVKPVCLPFGMLQPEQLCWISGWATEBK-GKTSEVLNAKVLIIET 407
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 352 TRCNADDAVQGEVTEKMKWACGIPGEGVDTCCGDSGGPLMYQSDQ-WHVVGIVSWGYGCGG 410
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 408 QRCNSRYVYDNLITPAMICAGFLQGVNDSQCGDSGGLVTSKNNIWLIGTWSGCGCAK 467
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 14
Q96T73 HUMAN
ID Q96T73 HUMAN PRELIMINARY; PRT; 492 AA.
AC Q96T73
DT 01-DEC-2001 (TREMELrel. 19, Created)
DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)
DT 01-WAR-2004 (TREMELrel. 26, Last annotation update)
DE Epitheliasin.
GN Name=TWPRSS2;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21223025; PubMed=11322890;
RA Jacquinet E., Rao N.V., Rao G.V., Wang Z., Albertine K.H.,
RA Hoidal J.R.;
RT "Cloning and characterization of the cDNA and gene for human
RT epitheliasin";
RL Eur. J. Biochem. 268:2687-2699(2001).
DR EMBL; AF329454; AAKS3559.1; -; mRNA.
DR HSP; P00766; 1CHG.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0005044; F:scavenger receptor activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001272; LDL receptor A.
DR InterPro; IPR001254; Peptidase S1_S6.
DR InterPro; IPR001314; Peptidase_S1a.
DR InterPro; IPR001190; S1r receptor.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00192; LDLa; 1.
DR SMART; SM00202; SR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS01209; LDLRA_1; 1.
DR PROSITE; PS00668; LDLRA_2; 1.
DR PROSITE; PS00287; SRCR_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
SQ SEQUENCE 492 AA; 53863 MW; 3ABA755BF276DADF CRC64;

Query Match 28.6%; Score 670.5; DB 2; Length 492;
Best Local Similarity 38.5%; Pred. No. 3.3e-48;
Matches 148; Conservative 59; Mismatches 128; Indels 49; Gaps 13;

Qy 75 CDGELDCPLGEDEHCHVCFPEGPAVAVRLSKDRSTLQVLDGATGNWFSACFDNFTFALA 134
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 133 CDGVSHCPGGEENRCVRLY--GP-----NFIQMYSSQKSWHPVCQDDWNENYG 181
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 135 ETACROMGYSKPTFRAVEIGPDQDLDDVVEITENSQELMRNMSG-----PCL 182
```

```
Db 182 RAACKOMGY--KNFYSSQ-----GIVDDSGTSFWMKLTNSAGNVDIYKLYHSDACS 232
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 183 SGLSVLSHCLACGKSL---KTPRVVVGGEASVDSWPQVSIQYDKQHVCGGSIIDPHWL 239
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 233 SKAVVSLRCIACGVNLSSRQSRIVGGESALPGAWPQVSLHVQNVHVCVCGSIIITPEWIV 292
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 240 TAAHCFRKH-TDVFVWVWVAGSDKLGSP-----PSLAVAKIIIIIEFNPMY---PKNDIAL 291
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 293 TAAHCFRKH-TDVFVWVWVAGSDKLGSP-----PSLAVAKIIIIIEFNPMY---PKNDIAL 348
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 292 MKLQPLTFSGTVRPICLPFDEBELTPATPLWILGWFTKONGKMSDILLOASVOVIDS 351
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 349 MKLQPLTFNDLVKPVCLPFGMLQPEQLCWISGWATEBK-GKTSEVLNAKVLIIET 407
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 352 TRCNADDAVQGEVTEKMKWACGIPGEGVDTCCGDSGGPLMYQSDQ-WHVVGIVSWGYGCGG 410
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 408 QRCNSRYVYDNLITPAMICAGFLQGVNDSQCGDSGGLVTSKNNIWLIGTWSGCGCAK 467
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 15
TMPS3 HUMAN
ID TMPS3 HUMAN STANDARD; PRT; 454 AA.
AC P5727; Q6ZMC3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Transmembrane protease, serine 3 (EC 3.4.21.-) (Serine protease TADG-
DE 12) (Tumor associated differentially-expressed gene-12 protein).
GN Name=TMPSR3; Synonyms=ECHOS1, TADG12; ORFNames=UNQ323/PRO382;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE (ISOFORMS A AND T).
RC TISSUE=Ovarian carcinoma;
RX PubMed=11068177; DOI=10.1016/S0925-4439(00)00058-2;
RA Underwood L.J., Shigemasa K., Tanimoto H., Beard J.B., Schneider E.N.,
RA Wang Y., Parmley T.H., O'Brien T.J.;
RT "Ovarian tumor cells express a novel multi-domain cell surface serine
RT protease.";
RL Biochim. Biophys. Acta 1502:337-350(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE (ISOFORMS A; B AND D), AND VARIANT ILE-53.
RX MEDLINE=20578749; PubMed=11137999; DOI=10.1038/83768;
RA Scott H.S., Kudoh J., Wattenhofer M., Shibuya K., Berry A., Chraast R.,
RA Guipponi M., Wang J., Kawasaki K., Asakawa S., Minoshima S.,
RA Younus F., Mehdi S.O., Radhakrishna U., Papasavvas M.P., Gehrig C.,
RA Rossier C., Korostishevsky M., Gal A., Shimizu N., Bonne-Tamir B.,
RA Antonarakis S.E.;
RT "Insertion of beta-satellite repeats identifies a transmembrane
RT protease causing both congenital and childhood onset autosomal
RT recessive deafness.";
RL Nat. Genet. 27:59-63(2001).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM E).
RX MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;
RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D.T., Brush J.,
RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,
RA Eaton D., Foster J.S., Grimaldi C., Gu Q., Hass P.E., Heidens S.,
RA Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,
RA Lewis L., Liao D., Mark M.R., Robbie E., Sanchez C., Schoenfeld J.,
RA Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,
RA Vandlen R.L., Watanabe C., Wian D., Woods K., Xie M.-H.,
RA Yansura D.G., Yu S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A.D.,
RA Wood W.I., Godowski P.J., Gray A.M.;
RT "The secreted protein discovery initiative (SPDI), a large-scale
```

RT effort to identify novel human secreted and transmembrane proteins: a
 RT bioinformatics assessment.";
 RL Genome Res. 13:2285-2270(2003).
 RN [4]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM E).
 RX PubMed=14702039; DOI=10.1038/ngl1285;
 RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
 RA Wakamatsu A., Hayashi K., Sato H., Negai K., Kimura K., Makita H.,
 RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
 RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
 RA Nagahori K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
 RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
 RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
 RA Omura Y., Abe K., Kamiyama M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
 RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
 RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hirakawa S., Chiba Y.,
 RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,
 RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
 RA Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M.,
 RA Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotseuka S.,
 RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
 RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
 RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
 RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
 RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
 RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
 RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
 RT "Complete sequencing and characterization of 21,243 full-length human
 RT cDNAs.";
 RL Nat. Genet. 36:40-45(2004).
 RN [5]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORMS A AND E).
 RX PubMed=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gurnatne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Faney J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [6]
 RP SUBCELLULAR LOCATION AND FUNCTION IN ENAC CLEAVAGE.
 RX MEDLINE=22281255; PubMed=12393794; DOI=10.1093/hmg/11.23.2829;
 RA Guipponi M., Vuesnigaux G., Wattenhofer M., Shibuya K., Vazquez M.,
 RA Doughterty L., Scamuffa N., Guida E., Okui M., Rossier C., Hancock M.,
 RA Buchet K., Raymond A., Hummel E., Marzella P.L., Kudoh J.,
 RA Shimizu N., Scott H.S., Antonarakis S.E., Rossier B.C.;
 RT "The transmembrane serine protease (TMPRSS3) mutated in deafness
 RT DFNB8/10 activates the epithelial sodium channel (ENAC) in vitro.";
 RL Hum. Mol. Genet. 11:2829-2836(2002).
 RN [7]
 RP VARIANTS DFNB8/DFNB10 CYS-251 AND LEU-404.

RE MEDLINE=21354482; PubMed=11462234; DOI=10.1002/humu.1159;
 RA Masmoudi S., Antonarakis S.E., Schwede T., Ghorbel A.M., Gratri M.,
 RA Pappaavas M.P., Driira M., Elgated-Bouilla A., Wattenhofer M.,
 RA Rossier C., Scott H.S., Ayadi H., Guipponi M.;
 RT "Novel missense mutations of TMPRSS3 in two consanguineous Tunisian
 RT families with non-syndromic autosomal recessive deafness.";
 RL Hum. Mutat. 18:101-108(2001).
 RN [8]
 RP VARIANTS DFNB8/DFNB10 TRP-109; PHE-194 AND ARG-407, AND VARIANTS
 RP ILR-53; SER-111 AND VAL-253.
 RX MEDLINE=21317610; PubMed=11494922; DOI=10.1136/jmg.38.6.396;
 RA Ben-Yosef T., Wattenhofer M., Riazuddin S., Ahmed Z.M., Scott H.S.,
 RA Kudoh J., Shibuya K., Antonarakis S.E., Bonne-Tamir B.,
 RA Radhakrishna U., Naz S., Ahmed Z., Riazuddin S., Pandya A.,
 RA Nance W.E., Wilcox E.R., Friedman T.B., Morell R.J.;
 RT "Novel mutations of TMPRSS3 in four DFNB8/10 families segregating
 RT congenital autosomal recessive deafness.";
 RL J. Med. Genet. 38:396-400(2001).
 RN [9]
 RP VARIANT DFNB8/DFNB10 GLY-103, AND VARIANTS ASN-173 AND THR-426
 RX MEDLINE=21904597; PubMed=11907649; DOI=10.1007/s00109-001-0310-6;
 RA Wattenhofer M., Di Iorio V., Rabionet R., Dougherty L., Pampalos A.,
 RA Schwede T., Montserrat-Sentis B., Arbones L., Iliades T.,
 RA Pasquadi-Biscaglia A., D'Amelio M., Alwan S., Rossier C., Dahl H.-H.M.,
 RA Petersen M.B., Estivill X., Gasparini P., Scott H.S.,
 RA Antonarakis S.E.;
 RT "Mutations in the TMPRSS3 gene are a rare cause of childhood
 RT nonsyndromic deafness in Caucasian patients.";
 RL J. Mol. Med. 80:124-131(2002).
 CC -!- FUNCTION: Probable protease. Seems to be capable of activating
 CC ENAC.
 CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Endoplasmic
 CC reticulum.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=5;
 CC Name=A;
 CC Name=B; Synonyms=C;
 CC Name=D; Synonyms=VSP_005391;
 CC Name=E; Synonyms=VSP_005392;
 CC Name=F; Synonyms=Truncated, TADG-12V;
 CC Name=G; Synonyms=VSP_005393, VSP_005394;
 CC Name=H; Synonyms=VSP_013184;
 CC -!- TISSUE SPECIFICITY: Expressed in many tissues including fetal
 CC cochlea. Isoform T is found at increased levels in some
 CC carcinomas.
 CC -!- PTM: Undergoes autolytic activation.
 CC -!- DISEASE: Defects in TMPRSS3 are a cause of childhood-onset
 CC autosomal recessive neurosensory deafness 8 (DFNB8) [MIM:601072].
 CC -!- DISEASE: Defects in TMPRSS3 are a cause of congenital autosomal
 CC recessive neurosensory deafness 10 (DFNB10) [MIM:605316].
 CC -!- SIMILARITY: Belongs to the peptidase S1 family.
 CC -!- SIMILARITY: Contains 1 LDL-receptor class A domain.
 CC -!- SIMILARITY: Contains 1 peptidase S1 domain.
 CC -!- SIMILARITY: Contains 1 SRCR domain.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL: AF201380; AAG37012.1; -; mRNA.
 CC EMBL: AB038157; BAB20077.1; -; mRNA.
 CC EMBL: AB038158; BAB20078.1; -; mRNA.
 CC EMBL: AB038159; BAB20079.1; -; mRNA.
 CC EMBL: AB038160; BAB20080.1; -; mRNA.
 CC EMBL: AK172842; BAB18806.1; -; mRNA.
 CC EMBL: AY359458; AAO88823.1; -; mRNA.
 CC EMBL: BC074846; AAH74846.1; -; mRNA.
 CC EMBL: BC074847; AAH74847.1; -; mRNA.

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 6, 2005, 19:48:25 ; Search time 5647 Seconds
(without alignments)
4378.768 Million cell updates/sec

Title: US-10-803-530-2
Perfect score: 2342
Sequence: 1 MDPSPDQLNSLDVPLRKP.....VTKVSALYINWVWRAEL 435

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2_1/USPTO_spool_p/US10803530/runat_05122005_083215_2119/app_query.fasta_1.583
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10803530 @CGN 1 1 7415 @runat_05122005_083215_2119 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

ALIGNMENTS

RESULT 1
AX076190
LOCUS AX076190 AX076190 1305 bp DNA linear PAT 06-FEB-2001
DEFINITION Sequence 1 from Patent WO0104141.
ACCESSION AX076190
VERSION AX076190.1 GI:12710815
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1
AUTHORS Suendermann, B., Hofmann, U., Matzku, S. and Wilbert, O.
TITLE Seripancrin
JOURNAL Patent: WO 0104141-A 1 18-JAN-2001;
MERCK PATENT GmbH (DE)

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Query Match	Length DB ID	Description
1	2342	100.0	1305 6	AX076190 Sequence
2	2338	99.8	2038 6	ARI142620 Sequence
3	2338	99.8	2038 6	BD137129 Human pro

4	2338	99.8	1314	6	AR5637419	Sequence
5	2337	99.8	1314	6	CQ924452	Sequence
6	2337	99.8	2021	8	AK172766	Homo sapi
7	2337	99.8	2079	8	BC011703	Homo sapi
8	2337	99.8	2081	6	AR454680	Sequence
9	2337	99.8	2081	6	AF179224	Homo sapi
10	2335	99.7	2121	6	AR559487	Sequence
11	2333	99.6	2590	6	CQ869992	Sequence
12	2329	99.4	1479	6	AX076192	Sequence
13	2328	99.4	2627	6	CQ869994	Sequence
14	2324	99.2	2070	6	BD235831	A novel m
15	2324	99.2	2079	6	AR232520	Sequence
16	2324	99.2	2079	6	AR454679	Sequence
17	2324	99.2	2079	6	AX207967	Sequence
18	2324	99.2	2079	8	AF216312	Homo sapi
19	2319	99.0	2137	6	AX207899	Sequence
20	2297.5	98.1	2063	6	CQ875679	Sequence
21	2297.5	98.1	2063	6	CS112264	Sequence
22	2297.5	98.1	2063	6	AX092380	Sequence
23	2297.5	98.1	2063	6	AX376262	Sequence
24	2297.5	98.1	2063	6	AX395214	Sequence
25	2297.5	98.1	2063	6	AX697206	Sequence
26	2297.5	98.1	2063	8	AY358530	Homo sapi
27	2167	92.5	1784	8	BC012752	Homo sapi
28	2044	87.3	1973	8	BC004855	Homo sapi
29	1844	78.7	2060	9	BC021368	Mus muscu
30	1844	78.7	2268	9	AY043240	Mus muscu
31	1782	76.1	2229	6	CQ869989	Sequence
32	1546.5	66.0	955	6	CQ729939	Sequence
33	1287	55.0	699	6	E13203	Human cDNA
34	1259	53.8	1189	6	AR559493	Sequence
35	860	36.7	2947	5	CR855611	Xenopus t
36	806	34.4	1008	6	BD194909	86 human
37	806	34.4	1008	6	CQ855237	Sequence
38	704.5	30.1	2992	9	BC042878	Mus muscu
39	699.5	29.9	2067	6	CQ875313	Sequence
40	699.5	29.9	2067	6	AX149581	Sequence
41	699.5	29.9	3320	9	BC085323	Mus muscu
42	698.5	29.8	1434	6	AR583622	Sequence
43	697.5	29.8	1314	6	AX190543	Sequence
44	697.5	29.8	1314	6	AX190546	Sequence
45	697.5	29.8	1689	6	CS075405	Sequence

/translation="MDPDSQQLNSLDVPLKRPIMETFRKVGPIPIIIALLSLASI
IIVVLKILDKYPLCOPLHFPKQLCDGELDCPLGEDEEHCVSFPEGPVAV
RLSKSTLQVLDLSATGNWFSACFDNFTALAEACROMGYSSKPTFRAVEIGPDOL
DVEITENSQELRMNSGCLSGSLVSLHCLACSKLTPRVVGEASVDSWPMOV
SIQYDKQHVCGGSLILDPHVLTAHCFKHTDVFNMKVRAGSKDLGSPSLAVAKII
IFENPMYKNDIALMKLQFPLTFSGTVRPICLPFDFBELTPTFLWIGWFTKQNG
GKMSDILLQASVQVVDSTRCNADDAVQSEVTEKMMCAPIPEGVDVTCQDGGGPLYQ
SDQHVVGIVSGVGGGSPSTPGVTVKYSVNLNWIYVNWKAEL"

ORIGIN

Alignment Scores:

Pred. No.: 3, 89e-208 Length: 1305
Score: 2342.00 Matches: 435
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-803-530-2 (1-435) x AX076190 (1-1305)

QY 1 MetAspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysPro 20
DB 1 ATGGATCTCTGACAGTATCACTCTGAACAGCCTCGATGTCAAACCCCTCGCAACCC 60
QY 21 ArgileProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSer 40
DB 61 CGTATCCCCATGGAGACCTTCAGAAAGGTGGGATCCCCATCATAGCACTACTGAGC 120
QY 41 LeuAlaSerIleIleIleValValLeuIleLysValIleLeuAspLysTyrPhe 60
DB 121 CTGGCGAGTATCATATTGTGTCTCATCAAGGTGATCTGGATAAATACTACTTC 180
QY 61 LeuCysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAsp 80
DB 181 CTCTCGCGGAGCCTCTCCACTTCATCCGAGGAGCAGCTGTGTGACGGAGAGCTGGAC 240
QY 81 CysProLeuGlyGluAspGluHisCysValLysSerPheProGluGlyProAlaVal 100
DB 241 TGTCCCTTGGGGGAGGAGCAGAGACACTGTGTCAAGAGCTTCCCGAAGGCGCTGCAAGTG 300
QY 101 AlaValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsn 120
DB 301 CGAGTCCGCTCTCCAGGACCGATCCACACTGACAGGTGTGGACTCGGCCACAGGGAAC 360
QY 121 TrpPheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGln 140
DB 361 TGGTCTCTGCTGTTTCGACAACTTCACAGAACTCTCGCTGAGACAGCTGTAGSCAG 420
QY 141 MetGlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeu 160
DB 421 ATGGCTACAGCAGCAAAACCCACTTTCAGAGCTGTGAGATTGGCCCGACAGAGGATCTG 480
QY 161 AspValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyPro 180
DB 481 GATGTGTGTGAATTCACAGAAAACAGCCAGGAGCTTCGCATCGGAACTCAAGTGGGCC 540
QY 181 CysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThr 200
DB 541 TGTCTCTCAGGCTCCCTTGGTCTCCCTGCATGTCTTGCCTGTGGGAGAGCCTGGAAGACC 600
QY 201 ProArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIle 220
DB 601 CCCCGTGTGTGGTGGGAGGAGGCTCTGTGGATTCTTGGCTTGGCGAGGTGACATC 660
QY 221 GluTyrAspLysGlnHisValCysGlySerIleLeuAspProHisTrpValLeuThr 240
DB 661 CAGTACACAAAACAGCAGCTCTGTGGAGGAGGATCCTGGACCCCACTGGGTCTCACG 720
QY 241 AlaAlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySer 260
DB 721 GCAGCCCACTGCTTTCAGAAAACATACCGATGTGTTCAACTGGAAAGGTGGGAGGCTCA 780
QY 261 AspLysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleIleGluPheAsn 280

DB 781 GACAACTGGGCAGCTCCCATCCCTGGCTGTGCCAAGATCATCATATTGAATTCAC 840
QY 281 ProMetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPhe 300
DB 841 CCCATGTATCCCAAGACAATGACATCGCCTCATGAAGCTGCAGTTCCCACTCCTTC 900
QY 301 SerGlyThrValArgProIleCysLeuProPhePheAspGluLeuThrProAlaThr 320
DB 901 TCAGGCACAGTCAGGCCCATCTGTCTGCCCTTCTTTCATGAGGAGCTCACTCCAGCCACC 960
QY 321 ProLeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIle 340
DB 961 CCACTCTGGATCATTTGATGGGCTTTACGAGCAGAAATGGAGGAGAGATGTCTGACATA 1020
QY 341 LeuLeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyr 360
DB 1021 CTGCTGCAGCGCTCAGTCCAGGTCAATTGACAGCACACGCTGCAATGCAGATGCGTAC 1080
QY 361 GlnGlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThr 380
DB 1081 CAGGGGAGTACCCGAGAAAGATGATGTGTGAGGAGATCCCGAAGGGGTGTGGACACC 1140
QY 381 CysGlnGlyAspSerGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGly 400
DB 1141 TGGCAGGGTGCACAGTGTGGGCGCTTGTGTGAGGAGATCCCGAAGGGGTGTGGACACC 1200
QY 401 IleValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValThrLysVal 420
DB 1201 ATCGTGTAGTGGGCTTATGGCTCGGGGCGCCGAGCACCCAGGAGTATACACCAAGTTC 1260
QY 421 SerAlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
DB 1261 TCAGCCTATCTCACTGGATCTACAATGTCTGGAGGCTGAGCTG 1305

RESULT 2

AR142620 2038 bp DNA linear PAT 08-AUG-2001
LOCUS Sequence 18 from patent US 6203979.
DEFINITION AR142620
ACCESSION AR142620
VERSION AR142620.1 GI:15103906
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2038)
AUTHORS Bandman, O., Hallman, J.L., Yue, H., Guegler, K.J., Corley, N.C.,
Tang, Y.Tom. and Shah, P.
TITLE Human protease molecules
JOURNAL Patent: US 6203979-A 18 20-MAR-2001;
FEATURES Location/Qualifiers
source 1..2038
/organism="unknown"
/mol_type="unassigned DNA"

Alignment Scores:

Pred. No.: 1, 6e-207 Length: 2038
Score: 2338.00 Matches: 434
Percent Similarity: 100.00% Conservatives: 1
Best Local Similarity: 99.77% Mismatches: 0
Query Match: 99.83% Indels: 0
DB: 6 Gaps: 0

US-10-803-530-2 (1-435) x AR142620 (1-2038)

QY 1 MetAspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysPro 20
DB 200 ATGGATCTCTGACAGTATCACTCTGAACAGCCTCGATGTCAAACCCCTCGCAACCC 259
QY 21 ArgileProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSer 40
DB 260 CGTATCCCCATGGAGACCTTCAGAAAGGTGGGATCCCCATCATAGCACTACTGAGC 319


```
Qy 41 LeuAlaSerIleIleValValValLeuIleLysValIleLeuAspLysTyrPhe 60
Db 320 CTGGCAGTATCATATTGTGTTCTCTCATCAAGGTGATTCTGGATAAATACTACTTC 379
Qy 61 LeuCysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAsp 80
Db 380 CTCTCGGGCAGCCTCTCCATCTCATCCGAGGAAGCAGCTGTGTGACGAGAGCTGGAC 439
Qy 81 CysProLeuGlyGluAspGluHisCysValLysSerPheProGluGlyProAlaVal 100
Db 440 TGTCTCTGGGGAGGAGCAGGAGCATTGTGTCAAGACTTCCCCGAAGGGCTGCAGTG 499
Qy 101 AlaValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsn 120
Db 500 GCAGTCCGCCCTCTCCAAGGACCGATCCACACTGCAGAGTGTGAGCTCGGCCACAGGGAAC 559
Qy 121 TrpPheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGln 140
Db 560 TGGTTCTCTGCTGTTTTCGACAACATTCACAGAAGCTCTCGCTGACAGAGCCTGTGAGGAG 619
Qy 141 MetGlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeu 160
Db 620 ATGGGCTACAGCAGCAAAACCCACTTTCAGAGCTGTGGAGATTGGCCCGACAGCAGGATCTG 679
Qy 161 AspValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyPro 180
Db 680 GATGTTGTTGAATCAGAGAAACAGCAGGAGCTTCGATGCGGAATCAAGTGGGCC 739
Qy 181 CysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThr 200
Db 740 TGTCTCTCAGGCTCCCTGCTCTCCCTGCACATGCTTCTGCTGTGGGAGAGCCTGAAGACC 799
Qy 201 ProArgValValGlyGlyGluAlaSerValAspSerTrpProTrpGlnValSerIle 220
Db 800 CCCCCTGTGTGGTGGGAGGAGGCTCTGTGGATTCTTGGCTTGGCAGGTTCAGCATC 859
Qy 221 GlnTyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThr 240
Db 860 CAGTACGACAAACACGACGCTGTGGAGGAGCATCTCGGACCCCACTGGGTCTCTCAGC 919
Qy 241 AlaAlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySer 260
Db 920 GCAGCCCACTGCTTCAGGAACATACCGATGTGTTCACTGGAAGGTGGGGCAGGCTCA 979
Qy 261 AspLysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsn 280
Db 980 GACAAACTGGGCAGCTTCCATCCCTGGCTGTGGCCAGATCATCATCAATTGAATCAAC 1039
Qy 281 ProMetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPhe 300
Db 1040 CCCATGTACCCCAAGACAATGACATCGCCCTCATGAAGCTGAGTTCCTCACTCACTTTC 1099
Qy 301 SerGlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThr 320
Db 1100 TCAGGCACAGTCAGGCCCATCTGTCTGCCCTTCTTTGATGAGGAGCTCACTCAGCCACC 1159
Qy 321 ProLeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIle 340
Db 1160 CCACTCTGGATCATTTGGATGGGGCTTTAGCAGCAGAGATGGAGGAGAGATGTCTGACATA 1219
Qy 341 LeuLeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyr 360
Db 1220 CTGCTGCAGCGTCACTGTCAGGTCAATTGACAGCACACGGGTGCAATGACAGCAGTGGTAC 1279
Qy 361 GlnGlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThr 380
Db 1280 CAGGGGGAAGTCACCGAGAAGATGATGTGTGAGGAGCATCCCGAAGGGGTGTGGACACC 1339
Qy 381 CysGlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGly 400
Db 1340 TGCCAGGGGTGACAGTGGTGGGGCCCTGTGATCAATCTGACACAGTGGCATGTGGTGGGC 1399
```

```
Qy 401 IleValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysVal 420
Db 1400 ATCGTTAGCTGGGCTATGGCTGGGGGGCCGAGCACCCAGGAGTATACACCAAGGTC 1459
Qy 421 SerAlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db 1460 TCAGCTATCTCAACTGGATCTACAATGTCTGGAAAGGCTGAGCTG 1504
RESULT 3
LOCUS BD137129 Human protease molecule. 2038 bp DNA linear PAT 18-SEP-2002
DEFINITION BD137129
ACCESSION BD137129
VERSION BD137129.1 GI:23232074
KEYWORDS JP 2002508970-A/6.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE 1 (bases 1 to 2038)
AUTHORS Bandman,O., Hillman,J.L., Yue,H., Guegler,K.J., Corley,N.C.,
Tang,T.Y. and Shah,P.
Human protease molecule
Patent: JP 2002508970-A 6 26-MAR-2002;
INCYTE PHARMACEUTICALS INC
OS Homo sapiens (human)
PN JP 2002508970-A/6
PD 26-MAR-2002
PF 12-JAN-1999 JP 2000540252
PI 16-JAN-1998 US 09/008271
PI OLGA BANDMAN,JENNIFER L HILLMAN,HENRY YUE,KARL J GUEGLER,NEIL
C CORLEY.
PI TOM Y TANG,PURVI SHAH
PC C12N15/09,A61K38/46,C07K16/40,C12N1/19,C12N1/21,C12N5/10 PC
,C12N9/48,C12N9/64,
PC C12Q1/68,C12N15/00,A61K37/54,C12N5/00
CC 1337018
FH Key Location/Qualifiers
FT source 1..2038
FT Location/Qualifiers
/organism='Homo sapiens (human)'
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'
ORIGIN
Alignment Scores: 1.6e-207 Length: 2038
Pred. No.: 2338.00 Matches: 434
Score: 2338.00 Conservative: 1
Percent Similarity: 100.00%
Best Local Similarity: 99.77% Mismatches: 0
Query Match: 99.83% Indels: 0
DB: 6 Gaps: 0
US-10-803-530-2 (1-435) x BD137129 (1-2038)
Qy 1 MetAspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysPro 20
Db 200 ATGGATCTCTGACATGATCAACCTCTGAACAGCCTCGATGTCAACACCTCTGGCAACCC 259
Qy 21 ArgIleProMetGluThrPheArgLysValGlyIleProIleIleAlaLeuSer 40
Db 260 CGTATCCCCATGGAGACCTTCAGAAAGGTGGGATCCCCCATCATCATAGCACTACTGAGC 319
Qy 41 LeuAlaSerIleIleIleValValValLeuIleLysValIleLeuAspLysTyrPhe 60
Db 320 CTGGCAGTATCATATTGTGTTCTCTCATCAAGGTGATTCTGGATAAATACTACTTC 379
Qy 61 LeuCysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAsp 80
Db 380 CTCTCGGGCAGCCTCTCCACTTCTCACTCCCGAGGAAGCAGCTGTGTGACGAGAGCTGGAC 439
```

```
QY      81  CysProLeuGlyGluAspGluHisCysValLysSerPheProGluGlyProAlaVal 100
Db      440 TGTCCCTTGGGGAGGAGCAGAGACACTGTGTCAAGAGCTTCCCGAAGGCGCTGCAGTG 499

QY     101  AlaValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsn 120
Db      500 GCAGTCCCGCTCTCCNAGGACCGATCCACACTGCAGGTGCTGGACTCGGCCACAGGGAAAC 559

QY     121  TrpPheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGln 140
Db      560 TGGTTCCTCGCTGTTTCGACAACTTCACAGAGCTCTCGCTGAGACGCTGTAGGCAG 619

QY     141  MetGlyTyrSerSerLysProThrPheArgAlaValGluLeGlyProAspGlnAspLeu 160
Db      620 ATGGGCTACAGCAGCAAAACCCACTTTCAGAGCTGTGGAGATTGGCCACAGACGAGTCTG 679

QY     161  AspValValGluLeuThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyPro 180
Db      680 GATGTTGTGTAATCACAGAAACAGCCAGGAGCTTCGCATGCCGAACTCAAGTGGGCC 739

QY     181  CysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThr 200
Db      740 TGTCTCTCAGGCTCCCTGCTCTCCCTGCACGTCTTGCCTGTGGGAGAGCTCGAAGACC 799

QY     201  ProArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIle 220
Db      800 CCCCCTGTGTGGTGGGGAGAGGCGCTCTGTGGATTCTTGGCCTTGGCAGGTCAGCATC 859

QY     221  GlnTyrAspLysGlnHisValCysGlyCysSerIleLeuAspProHisTrpValLeuThr 240
Db      860 CAGTACGACAAACAGCAGCTGTGGAGGAGCATCTCTGGACCCCACTGGGTCCTCAGC 919

QY     241  AlaAlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySer 260
Db      920 GCAGCCCACTGCTTCAGAAACATACCATGCTGTTCACCTGGAAAGGTGCGGCGAGGTCA 979

QY     261  AspLysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsn 280
Db      980 GACAAACTGGGCAGCTTCCCATCCCTGCTGTGCCAAGATCATCATTTGAATTCAAC 1039

QY     281  ProMetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPhe 300
Db     1040 CCCATGTACCCCAAGACAATGACATCGCCCTCATGAAGCTGCAGTTCCTCCACTCATTTC 1099

QY     301  SerGlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThr 320
Db     1100 TCAGGCACAGTCAGGCCCATCTGTCTGCCCTTCTTTGATGAGGAGCTCACTCCAGCCACC 1159

QY     321  ProLeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIle 340
Db     1160 CCACTCTGGATCAATTGGATGGGCTTTACGAAGCAGAAATGGAGGAAGATGCTTGACATA 1219

QY     341  LeuLeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyr 360
Db     1220 CTGCTGAGGGGTGAGTCCAGGTCAATGACAGCACACGGTGCATTCAGACGATGCGGTAC 1279

QY     361  GlnGlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThr 380
Db     1280 CAGGGGAGNATCAGCGAGAGATGATGTGCGAGCATCCCGAAGGGGGTGTGGACACC 1339

QY     381  CysGlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGly 400
Db     1340 TGCCAGGGGTGACAGTGGTGGGCCCTGATGTATCCAACTTGACCAAGTGGCATGTGGTGGC 1399

QY     401  IleValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysVal 420
Db     1400 ATCGTAGCTGGGCTATGGCTGCGGGGGCCGAGCACCCAGAGGATATACCAAGGTC 1459

QY     421  SerAlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db     1460 TCAGCCTATCTCAATGGATCTCAATGTCTGGAAAGGCTGAGCTG 1504
```

```
RESULT 4
AR637419      2038 bp      DNA      linear      PAT 20-APR-2005
LOCUS      Sequence 18 from patent US 6855811.
DEFINITION
AR637419      AR637419.1 GI:62771139
ACCESSION
AR637419.1
VERSION
AR637419.1
KEYWORDS
SOURCE
ORGANISM
Unknown.
REFERENCE
1 (bases 1 to 2038)
Bandman,O., Hillman,J.L., Yue,H., Guegler,K.J., Corley,N.C.,
Tang,Y.T. and Shah,P.
Human protease molecules
Patent: US 6855811-A 18 15-FEB-2005;
Incyte Pharmaceuticals, Inc.; Palo Alto, CA
FEATURES
source
1..2038
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 1.6e-207 Length: 2038
Score: 2338.00 Matches: 434
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.77% Mismatches: 0
Query Match: 99.83% Indels: 0
DB: 6 Gaps: 0
US-10-803-530-2 (1-435) x AR637419 (1-2038)
QY      1  MetAspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysPro 20
Db      200 ATGGATCTCTGACAGTGTATCAACTCTGACAGAGCTCTGATGTCAAAACCCCTGCGCAAAACC 259

QY     21  ArgIleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSer 40
Db      260 CGTATCCCATGGAGACCTTCAGAAAGGTGGGATCCCCATCATAGCACTACTAGC 319

QY     41  LeuAlaSerIleIleIleValValValLeuIleLysValIleLeuAspLysTyrPhe 60
Db      320 CTGGCAGATATCATCATTTGGTGTCTCATCAAGGTGATTTGGATAAATACTACTTC 379

QY     61  LeuCysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAsp 80
Db      380 CTCTGGGCGGAGCTCTCCACTTCATCCGAGGAAGCAGCTGTGTGACGAGAGCTGGAC 439

QY     81  CysProLeuGlyGluAspGluHisCysValLysSerPheProGluGlyProAlaVal 100
Db      440 TGTCCCTTGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTCCCGAAGGCGCTGCAGTG 499

QY     101  AlaValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsn 120
Db      500 GCAGTCCCGCTCTTCCAAGGACCGATCCACACTGCAGGTGCTGAGCTCGGCCACAGGGAAC 559

QY     121  TrpPheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGln 140
Db      560 TGGTTCCTCGCTGTTTCGACAACTTCAGAAAGCTCTCGCTGAGACGAGCTGTAGGCAG 619

QY     141  MetGlyTyrSerSerLysProThrPheArgAlaValGluLeGlyProAspGlnAspLeu 160
Db      620 ATGGGCTACAGCAGCAAAACCCACTTTCAGAGCTGTGGAGATTGGCCACAGACGAGTCTG 679

QY     161  AspValValGluLeuThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyPro 180
Db      680 GATGTTGTGTAATCACAGAAACAGCCAGGAGCTTCGCATGCCGAACTCAAGTGGGCC 739

QY     181  CysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThr 200
Db      740 TGTCTCTCAGGCTCCCTGCTCTCCCTGCACGTCTTGGCCTTGGCAGGTCAGCATC 799

QY     201  ProArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIle 220
```

```

Db      800  CCCCCTGTGGTGGGAGAGAGCCCTCTGTGGATTCTTGCCCTTGGCAGGTGACATC 859
Qy      221  GlnTyrAspLysGlnHisValCysGlySerIleLeuAspProHisTrpValLeuThr 240
Db      860  CAGTACGACAAACAGACAGCTGTGTGGAGGAGCATCTTGGACCCCACTGGGTCTCTCAG 919
Qy      241  AlaAlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgLysSer 260
Db      920  GCAGCCCACTGCTTTCAGGAAACATACCGATGTGTTCAACTGGAAGGTGCGGCGAGCTCA 979
Qy      261  AspLysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsn 280
Db      980  GACAACTGGGAGAGCTTCCCATCTGCTGTGGCCAGATCATCATATTGAATTCAC 1039
Qy      281  ProMetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPhe 300
Db      1040  CCCATGTACCCCAAGACAAATGACATCGCCCTCATGAAGTCGACGTCTCCACTCTTC 1099
Qy      301  SerGlyThrValArgProIleCysLeuProPheAspGluGluLeuThrProAlaThr 320
Db      1100  TCAGGCACAGTCAGGCCCATCTGTCTGCCCTTCTTTGATGAGGAGCTCACTCCAGCCACC 1159
Qy      321  ProLeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIle 340
Db      1160  CCACTCTGGATCATTTGGATGGGGCTTTACGAAGCAGAATGGAGGAAAGATGCTGACATA 1219
Qy      341  LeuLeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaTyr 360
Db      1220  CTGCTGCAGGCGTCAGTCCAGGTCATTGACAGCACACGCGTCAATGCAGACGATCGGTAC 1279
Qy      361  GlnGlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThr 380
Db      1280  CAGGGGGAAGTCACCGAGAAGATGATGTGTGCAGGCATCCCGAAGGGGGTGTGCACACC 1339
Qy      381  CysGlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValGly 400
Db      1340  TGCCAGGGTGACAGTGGTGGGGCCCTGTGATGACCAATCTGACCATGGCATGTGTGGGC 1399
Qy      401  IleValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysVal 420
Db      1400  ATCGTTAGTGGGCTATGGCTGCGGGGCGCCGAGCACCCAGGAGTATACACCAAGTTC 1459
Qy      421  SerAlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db      1460  TCAGCCTATCTCAACTGGATCTACAATGTCTGGAAGGCTGAGCTG 1504

RESULT 5
CQ924452      CQ924452      1314 bp      DNA      linear      PAT 23-NOV-2004
LOCUS         Sequence 1 from Patent WO2004097034.
DEFINITION   CQ924452
ACCESSION    CQ924452
VERSION      CQ924452.1  GI:56214165
KEYWORDS     Homo sapiens (human)
SOURCE       Homo sapiens
ORGANISM     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
              Hominoidea; Homo.
REFERENCE    1
AUTHORS      Golz,S., Brueggemeier,U., Geerts,A. and Polej,S.
TITLE        Diagnostics and therapeutics for diseases associated with human
              transmembrane serine protease 4 (tmprss4)
JOURNAL      Patent: WO 2004097034-A 1 11-NOV-2004;
              Bayer HealthCare AG (DE)
FEATURES     Location/Qualifiers
              source
              1..1314
              /organism="Homo sapiens"
              /mol_type="unassigned DNA"
              /db_xref="taxon:9606"

ORIGIN
Alignment Scores:

```

```

Pred. No.: 1,15e-207      Length: 1314
Score: 2337.00      Matches: 434
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 99.79%      Indels: 0
DB: 6      Gaps: 0

US-10-803-530-2 (1-435) x CQ924452 (1-1314)
Qy      2  AspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysProArg 21
Db      10  GATCTGACAGTGATCAACCTCTGAACAGCCTCGATGTCAAACCCCTCGCAACCCCTGT 69
Qy      22  IleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSerLeu 41
Db      70  ATCCCATGGAGACCTTCAGAAAGGTGGGATCCCATCATCATCATGACTACTGAGCCTG 129
Qy      42  AlaSerIleIleIleValValValLeuIleLysValIleLeuAspLysTyrTrpPheLeu 61
Db      130  GCGAGTATCATCATTTGTGTCTCATCAAGGTGATTCTGGATAAATACTACTTCTCTC 189
Qy      62  CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys 81
Db      190  TCGGGGAGCCTCTCCACTTATCCGAGGAGCAGCTGTGTGACGGAGAGCTGGACTGT 249
Qy      82  ProLeuGlyGluAspGluHisCysValLysSerPheProGluGlyProAlaValAla 101
Db      250  CCTTGGGGAGGACGAGGACCTGTGTCAAGAGCTTCCCGAAGGGCTCGAGTGGCA 309
Qy      102  ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp 121
Db      310  GTCCGCCCTCTCCAGGACCGATCCACTGCGAGGTGCTGGACTCGGCCACAGGGAACTGG 369
Qy      122  PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
Db      370  TTCTCTGCTGTTTCGACAACTTCACAGAAGCTCTCGCTGACAGCCTGTAGGCAGATG 429
Qy      142  GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
Db      430  GGCTACAGCAGCAAAACCCACTTTCAGAGCTGTGGAGATTGGCCAGACCAGGATCTGGAT 489
Qy      162  ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
Db      490  GTTGTGTAATACACAGAAACAGCAGGAGCTTCGCATGCGGAATCTCAAGTGGGCCCTGT 549
Qy      182  LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro 201
Db      550  CTCTCAGGCTCCCTGGTCTCCCTGCACCTGTCTTGGCTGTGGGAAGAGCTGGAAGACCCC 609
Qy      202  ArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
Db      610  CGTGTGGTGGGTGGGAGGAGCCCTCTGTGGATTCTTGGCTTGGCAGGTTCAGCATCCAG 669
Qy      222  TyrAspLysGlnHisValCysGlyClySerIleLeuAspProHisTrpValLeuThrAla 241
Db      670  TACGACAAACAGCAGCTGTGTGGAGGAGCATCTCTGGACCCCTGCTGGGTCTCTCAGGCA 729
Qy      242  AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
Db      730  GCCCCTGCTTCAGGAACATACCATGTGTTCAACTGGAAGGTGCGGCGAGGCTCAGAC 789
Qy      262  LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsnPro 281
Db      790  AAACCTGGCAGCTTCCCATCTGCTGGCTGTGGCCAAAGATCATCATCATGTAATTCACCCC 849
Qy      282  MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db      850  ATGTACCCCAAGACAAATGACATCGCCCTCATGAAGTGCAGTCTCCCACTCTTCTCA 909
Qy      302  GlyThrValArgProIleCysLeuProPheAspGluGluLeuThrProAlaThrPro 321
Db      910  GGCAAGTCAAGCCCATCTGTCTGCCCTTCTTTGATGAGGAGTCACTCCAGCCACCCCA 969

```


QY 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
 |||||
 Db 997 ATGTACCCCAAGACAATGACATGCCCTCATGAAGTGCAGTTCCTCCACTCATTCTCA 1056
 |||||
 QY 302 GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro 321
 |||||
 Db 1057 GGCACAGTCAGGCCCATCTGCTGCCCTCTTTGATGAGGAGCTCCTCCAGCCACCCA 1116
 |||||
 QY 322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu 341
 |||||
 Db 1117 CTCTGGATCTGGATGGGGCTTTACGAAGCAGAAATGAGGGAAGATGCTGACATACTG 1176
 |||||
 QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaTyrGln 361
 |||||
 Db 1177 CTCGAGGCGTCAGTCAGGTCATTGACAGCACACCGTGCAATGCGAGACGATGCGTACCAG 1236
 |||||
 QY 362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys 381
 |||||
 Db 1237 GGGGAAGTCACCGAAGATGATGTGAGGCAATCCCGAAGGGGGTGTGACACCTGC 1296
 |||||
 QY 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle 401
 |||||
 Db 1297 CAGGCTGACAGTGGTGGGCCCTGATGTACCAATCTGACCACTGCGATGCGTGGGCATC 1356
 |||||
 QY 402 ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSer 421
 |||||
 Db 1357 GTTACTGGGGCTATGGCTGGCGGGGCGCGAGCACCCAGGAGTATACCAAGGCTCTCA 1416
 |||||
 QY 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
 |||||
 Db 1417 GCCTATCAACTGGATCTACAATGCTCTGGAAGGCTGAGCTG 1458
 |||||

RESULT 7

LOCUS

BC011703 Homo sapiens transmembrane protease, serine 4, transcript variant 1, mRNA (CDNA clone MGC:19490 IMAGE:3610695), complete cds.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

BC011703.2 GI:33991397
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 2079)
 Strausberg RL, Feingold EA, Grouse LH, Derge JG, Klausner RD, Collins FS, Wagner L, Shenmen CM, Schuler GD, Altschul SF, Zeeberg B, Buetow KH, Schaefer CF, Bhat NK, Hopkins RF, Jordan H, Moore T, Max SI, Wang J, Hsieh F, Diatchenko L, Marusina K, Farmer AA, Rubin GM, Hong L, Stapleton M, Soares MB, Bonaldo MF, Casavant TL, Scheetz TE, Scapleton MJ, Udoin TB, Toshiyuki S, Carninci P, Prange C, Raha SS, Loquellano NA, Peters GJ, Abramson RD, Mullahy RJ, Bosak SA, McEwan PJ, McKernan KJ, Malek JA, Gunaratne PH, Richards S, Worley KC, Hale S, Garcia AM, Gay LJ, Hulyk SW, Villalón DK, Muzny DM, Sodergren EJ, Lu X, Gibbs RA, Fahey J, Helton E, Kettelman M, Madan A, Rodriguez S, Sanchez A, Whiting M, Madan A, Young AC, Shevchenko Y, Bouffard GG, Blakesley RW, Touchman JW, Green EP, Dickson MC, Rodriguez AC, Grimwood J, Schmutz J, Myers RM, Butterfield YS, Krzywinski MI, Skalek U, Smalios DE, Schnerch A, Schein JE, Jones SJ and Marra MA.
 Mammalian Gene Collection Program Team
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

CONSTRM

TITLE

human and mouse cDNA sequences

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

Submitted (30-JUL-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

REMARK COMMENT

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 On Aug 20, 2003 this sequence version replaced gi:15079794.

Contact: MGC help desk

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

Sequencing by: National Institutes of Health Intramural

Sequencing Center (NISC),

Gaithersburg, Maryland;

Web site: <http://www.nisc.nih.gov/>

Contact: nisc.mgc@nih.gov

Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,

Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,

Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,

Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,

Maduro, O.L., Magiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,

McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,

Tourgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,

Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAL Plate: 27 Row: f Column: 19

This clone was selected for full length sequencing because it

passed the following selection criteria: matched mRNA gi: 34304348.

FEATURES

Location/Qualifiers

1..2079

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:19490 IMAGE:3610695"

/tissue type="Pancreas, adenocarcinoma"

/clone_lib="NIH MGC 39"

/lab_host="DH10B-R"

/note="Vector: pOT7"

1..2079

/gene="TMPRSS4"

/note="synonyms: TMPRSS3, MT-SP2"

/db_xref="GeneID:56649"

/db_xref="MIM:606565"

208..1521

/gene="TMPRSS4"

/codon_start=1

/product="transmembrane protease, serine 4, isoform 1"

/protein_id="AAH11703.1"

/db_xref="GI:15079795"

/db_xref="GeneID:56649"

/db_xref="MIM:606565"

/translation="MLQDPDSQPLNSLDVKPLRPKPRIPMETFRKVGIPITIIALLSLA

SIIVVLIKVILOKVFYFLCQPLHIFIRKOLCDGLDPLGDEDEHCVKSPPEGPVAV

AVRLSKRSTLQVDSATGNWFSACFDNFTALAEACRMGYSSKPTFRAVEIGPDQ

DLDVVEITENSQELRMNNSGFCLSGLSVSLHCLACGSLKTPRVVGGEEASVDSKWP

QVSQYDKQHVCGGSLDHPHWLITAAHCRKHTDVFNKWRKVSGLKSPVGGEEASVDSKWP

IIIEFNPMYKNDIALMKLQPLFTFTSGTVPICLPFFDELTPTPLWIIQWGTQK

NGKMSDILQASVQVIDSTCNADDAYQGEVTEROMCAGIEGGVDTCQSGSGGLPM

YQSDQHWVGVISVWGCGGPGSTPGVTGVSAYLNIWYVWKAEL"

ORIGIN

Alignment Scores:

Pred. No.:

2.03e-207

Score: 2337.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 99.79%

DB: Indels: 0

Gaps: 0

US-10-803-530-2 (1-435) x BC011703 (1-2079)

QY

2 AppProAspSerAspGlnProLeuAsnSerLeuAspVallyAspProLeuArglyProArg 21

|||||

Db

217 GATCTGACAGTGATCAACCTCTGAACAGCCTCGATGTCAAACCCCTGCGCAACCCCGT 276

|||||

QY 22 IleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuSerLeu 41
Db ATCCCCATGGAGACCTTCAGAAAGGTGGGATCCCATCATCATAGCACTACTAGACCTG 336

QY 42 AlaSerIleIleIleValValLeuIleIleValLeuAspLysTyrPheLeu 61
Db GCGAGTATCATATGTGGTTGCTCATCAAGGTGATTCGTGATAAATACTACTTCTC 396

QY 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys 81
Db TCGGGGAGAGCTCTCCACTTCATCCCGAGAGCAGCTGTGTACCGAGAGCTGGAGCTG 456

QY 82 ProLeuGlyGluAspGluHisCysValLysSerPheProGluGlyProAlaValAla 101
Db CCCTTGGGGAGAGCAGAGGAGCACTGTGTCAAGAGCTTCCCGAAGGCGCTCAGTGGCA 516

QY 102 ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAenTrp 121
Db GTCCGCTCTCCAGGACCGATCCACACTGCAGGTGCTGGACTCGGCCACAGGGAAGTGG 576

QY 122 PheSerAlaCysPheAspAenPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
Db TTCTCTGCCTGTTCGACAACTTCACAGAAAGCTCTCGCTGACAGAGCTGTAGGAGATG 636

QY 142 GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
Db GGTACAGCAGCAAAACCCACTTTCAGAGCTGTGGAGATTGGCCAGACAGGATCTGGAT 696

QY 162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerClyProCys 181
Db GTTGTGTAATCACAGAAACAGCCAGAGAGCTTCGATCGGGAATCAAGTGGGCGCTGT 756

QY 182 LeuSerClySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro 201
Db CTCTCAGGCTCCTGGTCTCCCTGCAGCTCTTGTGCTGTGGAGAGCTTGAAGCCCC 816

QY 202 ArgValValGlyGlyGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
Db CGTGTGGTGGTGGGAGAGGCGCTCTGTGGATTCTTGGCCTTGGCAGGTGAGCATCCAG 876

QY 222 TyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAla 241
Db TAGCAAAACAGACAGTCTGTGGAGGAGCATCTGGACCCCACTGGTCTCTCAGCGCA 936

QY 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
Db GCCCACTGCTTCAGAAACATACCGATGTGTCAACTGGAAGGTGCGGCGAGGCTCAGAC 996

QY 262 LysLeuGlySerPheProSerLeuAlaValAlaValIleIleIleGluPheAsnPro 281
Db AAACCTGGGCGAGCTTCCATCCCTGGCTGTGGCCAGATCATCATTTGAATTCAACCCC 1056

QY 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db ATGTACCCCAAGACAAATGACATGCCCTCATGAAGCTGCAGTTCCCACTTCTCA 1116

QY 302 GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro 321
Db GGCACAGTCAGGCCCATCTGTCTGCCCTTCTTGTATGAGGAGTCACTCCAGGACAGCCCA 1176

QY 322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu 341
Db CTCTGGATCATTTGGTGGGCTTTACGAGCAGATGGAGGAGATGTCTGACATACTG 1236

QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaThrGln 361
Db CTGACGGCGTCAGTCCAGGTCATTGACAGACACGGTGTCAATGCAGACGATGGTACCAG 1296

QY 362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys 381
Db GGGGAAGTCACCGAGAAGATGATGTGTGACAGGATCCCGAAGGGGTGTGGACACCTGC 1356

QY 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle 401
Db CAGGGTGACAGTGGTGGCCCTGATGTACCAATCTGACCAGTGGCATGTGTGGGCATC 1416

QY 402 ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSer 421
Db GTTAGCTGGGCTATGGCTGCGGGGGCCGAGCAGCCAGGAGTATACCAAGGTCTCA 1476

QY 422 AlaTyrLeuAenTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db GCCTATCTCAACTGGATCTACAATGTCTGGAAGGCTGAGCTG 1518

RESULT 8
AR454680
LOCUS AR454680
DEFINITION Sequence 7 from patent US 6682890.
ACCESSION AR454680
VERSION AR454680.1 GI:42688149
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 2081)
AUTHORS Mack,D., Gish,K.C. and Wilson,K.E.
TITLE Methods of diagnosing and determining prognosis of colorectal cancer
JOURNAL Patent: US 6682890-A 7 27-JAN-2004;
Protein Design Labs, Inc.; Fremont, CA
FEATURES
source Location/Qualifiers
1..2081
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Alignment Scores:
Pred. No.: 2,04e-207 Length: 2081
Score: 2337.00 Matches: 434
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.79% Indels: 0
DB: 6 Gaps: 0

US-10-803-530-2 (1-435) x AR454680 (1-2081)

QY 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysProArg 21
Db GATCTGACAGTATCAACCTCTGAACAGCTCGATGTCAACCCCTGCGCAACCCCTG 283

QY 22 IleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSerLeu 41
Db ATCCCATGGAGACCTTCAGAAAGGTGGGATCCCATCATCATAGCACTACTGAGCCTG 343

QY 42 AlaSerIleIleIleValValLeuIleLysValIleLeuAspLysTyrPheLeu 61
Db GCGAGTATCATATGTGGTTGCTCATCAAGGTGATTCGATAAATACTACTTCTC 403

QY 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys 81
Db TCGGGGAGAGCTCTCCACTTCATCCCGAGAGAGAGCTGTGTGACGGAGAGCTGAGCTGT 463

QY 82 ProLeuGlyGluAspGluHisCysValLysSerPheProGluGlyProAlaValAla 101
Db CCCTTGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTCCCGAAGGCGCTGCGATGGCA 523

QY 102 ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp 121
Db GTCCGCTCTCCAGGACCGATCCACACTGCAGGTGTGGACTCGGCCACAGGGAATCG 583

QY 122 PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
Db TTCTCTGCCTTTCGACAACTTCACAGAAAGCTCTCGCTGAGCAGAGCTGTAGGACATG 643

QY 142 GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161

Db 644 GCCTACAGCAGCAAAACCCACTTTCAGAGCTGTGGAGATTGGCCCAAGACAGGATCTGGAT 703
Qy 162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
Db 704 GTTGTGTAATCACAGAAACAGCCAGGAGCTTCGCATCGCGAACTCAAGTGGCCCTGT 763
Qy 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlySerLeuHisThrPro 201
Db 764 CTCTCAGGCTCCCTGGTCTCCCTGCACATGTCTTGTGCTGTGGAAAGAGCTGAAGACCCCC 823
Qy 202 ArgValValGlyGlyGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
Db 824 CGTGTGTGGTGGGAGAGAGCCCTCTGTGGATTCTTGGCCCTTGACAGGTTCAGCATCCAG 883
Qy 222 TyrAspLysGlnHisValCysGlySerIleLeuAspProHisTrpValLeuThrAla 241
Db 884 TACGACAAACAGCACGCTCTGTGGAGGAGCATCTCTGGACCCCACTGGTCTCTCAGGCA 943
Qy 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
Db 944 GCCCACTGCTTCAGGAAACATACCGATGTGTTCAACTGGAAGGTGCGGCGAGCTCAGAC 1003
Qy 262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleIleGluPheAsnPro 281
Db 1004 AAATCTGGGAGCTTCCCATCCCTGGCTGTGGCAAGATCATCATTAATTCACACCCC 1063
Qy 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db 1064 ATGTACCCCAAGACAAATGACATCGCCCTCATGAAGCTGCAGTCTCCACTCACTTTCICA 1123
Qy 302 GlyThrValArgProIleCysLeuProPheAspGluGluLeuThrProAlaThrPro 321
Db 1124 GGCACAGTCAGGCCCATCTGTCTGCCCTTCTTTGTATGAGGAGCTCACTCCAGCCACCCA 1183
Qy 322 LeuTrpIleLeuGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu 341
Db 1184 CTCTGGATCATTGGATGGGCTTTACGAAGCAAGATGGAGGAAAGATGTCTGACATATCG 1243
Qy 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaThrGln 361
Db 1244 CTGCAGGCGTCAGTCCAGGTCAATGCAGCACACGGTGCATGCAGAGCATGCGTACCAG 1303
Qy 362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys 381
Db 1304 GGGGAAGTCACCGAAGATGATGTGTGAGGATCCCGAAGGGGTGTGGACACCTGC 1363
Qy 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle 401
Db 1364 CAGGCTGACAGTGGTGGGCCCCCTGATGTACCAATCTGACCACTGTGGTGGGCATC 1423
Qy 402 ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSer 421
Db 1424 GTTACTGGGCTATGGCTGCGGGGCCCCGAGCACCCAGGAGTATACCAAGGTCTTCA 1483
Qy 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db 1484 GCCTATCTCACTGGATCTACAATGTCTGGAAGGCTGAGCTG 1525

RESULT 9

AF179224

LOCUS

AF179224 Homo sapiens transmembrane serine protease 3 (TMPRSS3) mRNA, linear PRI 08-JUN-2000

DEFINITION

complete cds.

ACCESSION

AF179224

VERSION

AF179224.1

KEYWORDS

GI:8347148

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Homidae; Homo.

REFERENCE

1 (bases 1 to 2081)

AUTHORS

Wallrapp,C., Hahnel,S., Muller-Pillasch,F., Burghardt,B.,

Iwamura,T., Ruthenburger,M., Lerch,M.M., Adler,G. and Gress,T.M.

A novel transmembrane serine protease (TMPRSS3) overexpressed in

pancreatic cancer

Cancer Res. 60 (10), 2602-2606 (2000)

10825129

2 (bases 1 to 2081)

Wallrapp,C. and Gress,T.M.

Direct Submission

Submitted (20-AUG-1999)

Internal Medicine I, University of Ulm,

Robert-Koch-Street 8, Ulm, Baden-Wuerttemberg 89081, Germany

Location/Qualifiers

1..2081

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/chromosome="11"

/map="11q23.3"

/tissue_type="pancreatic carcinoma"

1..2081

/gene="TMPRSS3"

215..1528

/gene="TMPRSS3"

/codon_start=1

/product="transmembrane serine protease 3"

/protein_id="AAF74526.1"

/db_xref="GI:8347149"

/translation="MLQPPDSQPLNSLDVPLKRPRIEMETFRKVGIPITIALISLA

SIIVLVILKIVLYFLCQPLHFI PRKOLCDGELDCPLGEDHEHCVKSPPEGPVAV

AVRLSKORSTLQVLDSATGNWFSACFDNFTEALAEACROMGYSSKPTFRVAEIGPDQ

DLDVVEITENSQELRMNNSGFCLSGLVSLHCLACGKSLKTPTRVVGEEASVDSWPW

QVSIOYDKHQVCGGSILDPHVLWTAACHFRKHTDFVNWKVRAGSDKLGSPSLAVAKI

IIIFNPKNDIALMKLQPLTFSGTVRPICLPFFDELTPTATPLWIIWGGFTKQ

NGKMSDILLQASVQVLDSTRCNADDAVQGEVTEKMCAGIPEGGVDTCCQSDSGGPLM

YQSDQWVHVIVSWGCGGPGSPGTVYTKVSAYLNLWVYVWKAEL"

215..307

/gene="TMPRSS3"

/note="Region: cytoplasmic domain"

308..373

/gene="TMPRSS3"

/note="Region: transmembrane domain"

374..1525

/gene="TMPRSS3"

/note="Region: extracellular domain"

602..604

/gene="TMPRSS3"

/note="glycosylation site"

746..748

/gene="TMPRSS3"

/note="glycosylation site"

order(800..802,1142..1144)

/gene="TMPRSS3"

/note="disulfide bond"

order(902..904,950..952)

/gene="TMPRSS3"

/note="disulfide bond"

947..949

/gene="TMPRSS3"

/note="active site"

1082..1084

/gene="TMPRSS3"

/note="active site"

order(1280..1282,1328..1330)

/gene="TMPRSS3"

/note="disulfide bond"

1373..1375

/gene="TMPRSS3"

/note="active site"

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	99.79%	Indels:	0
DB:	8	Gaps:	0
US-10-803-530-2 (1-435) x AF179224 (1-2081)			
QY	2	AspProAspSerAspGlnProLeuAenSerLeuAenValLysProLeuArgLysProArg	21
Db	224	GATCTGACAGTGATCAACCTCTGACAGCCTCGATGTCAAACCCCTCGGCAACCCCGT	283
QY	22	IleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSerLeu	41
Db	284	ATCCCATGAGACCTTCAGAAAGTGGGATCCCATCATCATAGCACTACTGAGCCTG	343
QY	42	AlaSerIleIleIleValValValLeuLeuLeuValIleLeuApsLysTyTyPheLeu	61
Db	344	GGAGTATCATCATTTGGTGTCTCATCAAGGTGATTCGGATAAATACTACTTCCTC	403
QY	62	CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys	81
Db	404	TGCGGGAGCCTCTCCACTTCATCCGAGGAGCAGCTGTGTGACGGAGAGCTGGACTGT	463
QY	82	ProLeuGlyGluAspGluHisCysValLysSerPheProGluGlyProAlaValAla	101
Db	464	CCCTTGGGGAGGAGCAGGAGCACTGTGTCAAGAGCTTCCCGAAGGGCCCTGCAGTGGCA	523
QY	102	ValArgLeuSerLysAspArgSerThrLeuGlnValLeuApsSerAlaThrGlyAsnTrp	121
Db	524	GTCCGCTCTCCAGGACCGATCCACTGCTGAGGTGTGGACTCGGCCACAGGGAATGG	583
QY	122	PheSerAlaCysPheAspAsnPhetrGluAlaLeuAlaGluThrAlaCysArgGlnMet	141
Db	584	TTCTCTGCTGTTCGAACTTCACAGAGCTCTCGCTGAGACAGCTGTAGCAGATG	643
QY	142	GlyTrpSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp	161
Db	644	GGCTACAGCAGAAACCCACTTTCAGAGCTGTGAGATTTGCCCGACAGCAGGATCTGGAT	703
QY	162	ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys	181
Db	704	GTGTGTGAATCACAGAAACAGCAGGAGCTTTCGATCGGAACTCAAGTGGGCCCTGT	763
QY	182	LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro	201
Db	764	CTCTCAGGCTCCCTGGTCTCCCTGCACTGTCTTGCCTGTGGGAAGAGCCTGAAGACCCC	823
QY	202	ArgValValGlyGlyGluAlaSerValAspSerTrpProTrpGlnValSerIleGln	221
Db	824	CGTGTGTGGTGGGAGGAGGCTCTGTGGATTTCTTGGCCTTGGCAGGTCCAGCATCCAG	883
QY	222	TyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAla	241
Db	884	TACGACAAACAGCAGCTCTGTGGAGGAGCATCTGGACCCCCACTGGTCTCCACGGCA	943
QY	242	AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp	261
Db	944	CCCCACTGCTTCAGAAACATACCGATGTGTCAACTGGAAGGTGGCGGAGGCTCAGAC	1003
QY	262	LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsnPro	281
Db	1004	AACTGGGAGCTTCCCATCCCTGGCTGTGGCCAGATCATCATTTGAATTCACCCCC	1063
QY	282	MetTrpProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer	301
Db	1064	ATGATACCCCAAGACAAATGACATCGCCCTCATGAAGCTGAGTTCCTCCACTTCTCA	1123
QY	302	GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro	321
Db	1124	GGCAGCTCAGGCCCATCTGTCTGCCCTTCTTGTATGAGGAGCTCACTCCAGCACCCCCA	1183
QY	322	LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyLysMetSerAspIleLeu	341
Db			
Db	1184	CTCTGGATCATTTGGATGGGGCTTTTACGAAGACAGAAATGAGGAGATGTCTGCATACTG	1243
QY	342	LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTrpGln	361
Db	1244	CTGCAGGCGTCAGTCCAGGTTCATTGACAGCACACGGTGCATGACAGCATGGGTACCAG	1303
QY	362	GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys	381
Db	1304	GGGGAAGTCAACGAGAAGATGATGTGTGAGGATCCCGGAAGGGGTGTGGACACCTGC	1363
QY	382	GlnGlyAspSerGlyGlyProLeuMetTyTrpGlnSerAspGlnTrpHisValValGlyIle	401
Db	1364	CAGGTGACAGTGTGGGCCCTTGATGTACCAATCTGACAGTGGCATGTGGTGGGCATC	1423
QY	402	ValSerTrpGlyTyTrpGlyCysGlyGlyProSerThrProGlyValTyTrpLysValSer	421
Db	1424	GTTAGCTGGGCTATGGCTGCGGGGCCCGAGCACCCAGGAGTATACACCAAGGTCTCA	1483
QY	422	AlaTyTrpLeuAsnTrpIleTyTrpAsnValTrpLysAlaGluLeu	435
Db	1484	GCCTATCTCAACTGGATCTACAAATGTCTGGAAGGCTGAGCTG	1525
RESULT 10			
AR559487			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
ORIGIN			
Alignment Scores:			
Pred. No.:	3-2e-207	Length:	2121
Score:	2335.00	Matches:	434
Percent Similarity:	99.77%	Conservative:	0
Best Local Similarity:	99.77%	Mismatches:	1
Query Match:	99.70%	Indels:	0
DB:	6	Gaps:	0
US-10-803-530-2 (1-435) x AR559487 (1-2121)			
QY	1	MetAspProAspSerAspGlnProLeuAenSerLeuApsValLysProLeuArgLysPro	20
Db	277	ATGGATCTCTGACAGTATCAACCTCTGAACAGCCTCGATGTCAAACCCCTCGCAACCC	336
QY	21	ArgIleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSer	40
Db	337	CGTATCCCCATGGAGACCTTCAGAAAGTGGGATGCCCATCATCATAGCACTACTGAGC	396
QY	41	LeuAlaSerIleIleIleValValValLeuIleLysValIleLeuAspLysTyTrpPhe	60
Db	397	CTGGCAGTATCATCATTTGGTGTCTCTCATCAAGGTGATTTCTGGATAAATACTACTTC	456
QY	61	LeuCysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAsp	80
Db	457	CTCTGGGCGACCTCTCCACTTCATCCCGAAGAGCAGCTGTGTGACGGAGAGCTGGAC	516
QY	81	CysProLeuGlyGluAspGluHisCysValLysSerPheProGluGlyProAlaVal	100
Db	517	TGTCCCTTTGGGGAGGAGCAGGAGGACACTGTGTCAAGAGCTTCCCCGAAGGGCTGCAGTG	576


```
QY 101 AlaValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsn 120
Db 577 GCAGTCCGCTCTCCAAAGACCGATCCACACTGCAGGTCGTGGACTCGCCACAGGGAAC 636
QY 121 TrpPheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGln 140
Db 637 TGGTTCCTGCTGTTTCGACAACTTCAGAAAGCTCTCGTGAGACAGCCTGTAGGCAG 696
QY 141 MetGlyTyrSerSerLysProThrPheArgAlaValGluLeGlyProAspGlnAspLeu 160
Db 697 ATGGGCTACAGCAGCAAAACCCACTTTTCAGAGCTGTGGAGATGGGCCAGACCAGGATCTG 756
QY 161 AspValValGluLeuThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyPro 180
Db 757 GATGTTGTTGTAATCACAGAAACAGCCAGGAGCTTCGCATCGGAACTCAAGTGGGCC 816
QY 181 CysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThr 200
Db 817 TGTCTCTCAGGCTCCCTGCTCTCCCTGCACCTGCTTGGCTGTGGAAAGAGCCTGAAGACC 876
QY 201 ProArgValValGlyGlyGluAlaSerValAspSerTrpProTrpGlnValSerIle 220
Db 877 CCCCCTGTGGTGGGGAGAGGCCCTCTGTGGATTCTTGGCCCTTGGCAGGTACGATC 936
QY 221 GlnTyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThr 240
Db 937 CAGTACGACAAACAGCAGCTCTGTGGAGGAGCATCTCGACCCCACTGGTCTCAGC 996
QY 241 AlaAlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySer 260
Db 997 GCAGCCCACTGCTTCAGGAAACATACCCGATGTGTTCAACTGGAAGGTGGCGGAGCTCA 1056
QY 261 AspLysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsn 280
Db 1057 GACAACTGGGAGGTTCCTCACTCCCTGGCTGTGGCCAAAGATCATCATCAATTCAC 1116
QY 281 ProMetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPhe 300
Db 1117 CCCATGTACCCCAAGACATGACATCGCCCTCATGAAGCTGCAGTTCCTCACTTTC 1176
QY 301 SerGlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThr 320
Db 1177 TCAGGCACAGTCAGGCCCATCTGTCTGCCCTCTTTGTATGAGGAGCTCACTCAGCCACC 1236
QY 321 ProLeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyLysMetSerAspIle 340
Db 1237 CCACTCTGGATCATTTGGATGGGCTTTACGAAGCAGATGGAGGAAAGATGCTGCACATA 1296
QY 341 LeuLeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyr 360
Db 1297 CTGCTGCAGGCTCAGTCCAGGTCATTGACAGCACACCGGTGCAATGCAGACGATCGTAC 1356
QY 361 GlnGlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyValAspThr 380
Db 1357 CTGGGGGAAGTCACCGAGAAGATGATGTGTGCAGGCATCCCGAAAGGGGTGTGCACACC 1416
QY 381 CysGlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGly 400
Db 1417 TCCAGGGTGACAGTGGTGGGCCCTGATGTACCAATCTGACACGTGGCATGTGTGGTGGCC 1476
QY 401 IleValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysVal 420
Db 1477 ATCGTTAGCTGGGCTATGGCTCGGGGGCCCCGAGCACCCAGGGGTATACACCAAGTTC 1536
QY 421 SerAlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db 1537 TCAGCCTATCTCAACTGGATCTACAATGTCTGGAAGGCTGAGCTG 1581
```

RESULT 11

```
CQ869992
LOCUS          2590 bp      DNA      linear      PAT 13-SEP-2004
DEFINITION    Sequence 413 from Patent WO2004074320.
ACCESSION     CQ869992
```

```
VERSION        CQ869992.1      GI:51999709
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE      1
AUTHORS       Morris,D.W., Morris,D.W. and Malandro,M.S.
TITLE         Novel therapeutic targets in cancer
JOURNAL       Patent: WO 2004074320-A 413 02-SEP-2004;
              Sagres Discovery, Inc. (US)
FEATURES       Location/Qualifiers
               source
               1..2590
               /organism="Homo sapiens"
               /mol_type="unassigned DNA"
               /db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.:      6,32e-207      Length:      2590
Score:          2333.00        Matches:    434
Percent Similarity: 99.77%      Conservative: 0
Best Local Similarity: 99.77%    Mismatches: 1
Query Match:    99.62%          Indels:     0
Db:             6              Gaps:       0
US-10-803-530-2 (1-435) x CQ869992 (1-2590)
QY 1 MetAspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysPro 20
Db 226 ATGGATCTGACAGTGATCAACCTCTGAACAGCCTCGATGTCAAACCCCTCGCACAACCC 285
QY 21 ArgIleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSer 40
Db 286 CGTATCCCATCGAGACCTTCAGAAAGGTGGGATCCCATCATCATAGCACTACTGAGC 345
QY 41 LeuAlaSerIleIleIleValValValLeuIleLysValIleLeuAspLysTyrTyrPhe 60
Db 346 CTGGCGAGTATCATCATTTGTGTTCTCTCATCAAGGTGATTCTCGATAAATACTACTTTC 405
QY 61 LeuCysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAsp 80
Db 406 CTCTCGGGGAGCCTCTCCACTTCATCCGAGGAGCAGCTGTGTGACGGAGAGCTGGAC 465
QY 81 CysProLeuGlyGluAspGluHisCysValLysSerPheProGluGlyProAlaVal 100
Db 466 TGTCCCTTTGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTCCCGAAGGCGCTCGAGTG 525
QY 101 AlaValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsn 120
Db 526 GCAGTCCGCTCTCCAAAGACCGATCCACACTGCAGGTGCTGGACTCGGCCACAGGGAAC 585
QY 121 TrpPheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGln 140
Db 586 TGGTTCCTGCTGTTTCGACAACTTCAGAAAGCTCTCGCTGACAGACGCTGTAGGCAG 645
QY 141 MetGlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeu 160
Db 646 ATGGGCTACAGCAGCAAAACCCACTTTCAGAGCTGTGGAGATGGGCCACAGGATCTG 705
QY 161 AspValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyPro 180
Db 706 GATGTTGTTGAAATCACAGAAACAGCCAGGAGCTTCGCATCGGAACTCAAGTGGGCC 765
QY 181 CysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThr 200
Db 766 TGTCTCTCAGGCTCCCTGCTCTCCCTGCACCTGTCTTGTCTGTGGAAAGAGCCTGAAGAC 825
QY 201 ProArgValValGlyGlyGluAlaSerValAspSerTrpProTrpGlnValSerIle 220
Db 826 CCCCCTGTGGTGGGTGTGAGGAGGCCCTCTGTGGATTCTTGGCCTTGGCAGGTTCAGCATC 885
```

```
QY 221 GlnTyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTyrValLeuThr 240
|
Db 886 CAGTACGACAAACAGCACGTCTGTGGAGGAGCATCTCGGACCCCATCGGTCTCAG 945
|
QY 241 AlaAlaHisCysPheArgLysHisThrAspValPheAsnTyrPheValArgAlaGlySer 260
|
Db 946 GAGGCCCATCTGCTTCAGGAAACATACCGATGTGTTCACCTGGAAGGTGCGGCGAGCTCA 1005
|
QY 261 AspLysLeuGlySerPheProSerLeuAlaValAlaValIleIleIleIleIleIleIle 280
|
Db 1006 GACAACTGGGAGCTTCCATCTCTGGCTGTGGCCAGATCATCATTTGAATCAAC 1065
|
QY 281 ProMetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPhe 300
|
Db 1066 CCCATGTACCCCAAGACATGACATGCCCTCATGAAGCTGCATCTCCACTCATCTTC 1125
|
QY 301 SerGlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThr 320
|
Db 1126 TCAGGCACAGTCAGGCCCATCTGTCTGCCCTTCTTTGATGAGGAGCTCACTCCAGCCACC 1185
|
QY 321 ProLeuTyrPheIleGlyTyrGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIle 340
|
Db 1186 CCACCTCTCGATCATTTGGATGGGGCTTTACGAAGCAGAATGGAGGGAAGATGCTGACATA 1245
|
QY 341 LeuLeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaTyr 360
|
Db 1246 CTGCTCGAGGCTCAGTCCAGTCAATGACAGCACACCGTGCAATGCAGACATGCGTAC 1305
|
QY 361 GlnGlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThr 380
|
Db 1306 CAGGGGGAAGTCACCCAGAGAAGATGATGTGTCAGGCATCCCGAAGGGGTGTGACACC 1365
|
QY 381 CysGlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTyrPheValValGly 400
|
Db 1366 TCCAGGGTGACAGTGGTGGGGCCCTGATGTACCAATCTGACCAAGTGGCATGTGGTGGGC 1425
|
QY 401 IleValSerTyrGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysVal 420
|
Db 1426 ATCGTGTAGTGGGGCTATGGCTGCGGGGGCCGAGCACCCAGAGATATACCAAGTGC 1485
|
QY 421 SerAlaTyrLeuAsnTyrPheTyrAsnValTyrPheAlaGluLeu 435
|
Db 1486 TCAGCCTATCACTGATCTACAAATGTCTGGAAGGCTGAGCTG 1530
|
```

```
RESULT 12
AX076192
LOCUS AX076192 1479 bp DNA linear PAT 06-FEB-2001
DEFINITION Sequence 3 from Patent WO0104141.
ACCESSION AX076192
VERSION AX076192.1 GI:12710817
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1
REFERENCE
AUTHORS Suendermann, B., Hofmann, U., Matzku, S. and Wilbert, O.
TITLE Serpincrin
JOURNAL Patent: WO 0104141-A 3 18-JAN-2001;
MERCK PATENT GmbH (DE)
FEATURES
source
1..1479
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
1..1479
/feature="unnamed protein product"
/codon_start=1
/protein_id="C28480.1"
/db_xref="GI:12710818"
/translation="MDPDSQPLNSLDVFKLPKPRIPMETFRKVGIPPIIALLSLASI
IIVVLIKVLIDKYFLCGQPLHP IPRKQLCDGELDCLGDEBEHCVKSPFEGPAVA
```

CDS

source

FEATURES

JOURNAL

TITLE

AUTHORS

REFERENCE

SOURCE

KEYWORDS

VERSION

ACCESSION

DEFINITION

LOCUS

AX076192

RESULT 12

Db 781 GACAACTGGGCGAGCTTCCCATCCCTGGCTGTGGCCAAAGATCATCATTTGAATTCAC 840
Qy 281 ProMetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPhe 300
Db 841 CCCATGTACCCCAAGACAATGACATCGCCCTCATGAAGCTGCAGTTCCTCCACTCACTTTC 900
Qy 301 SerGlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThr 320
Db 901 TCAGGCACAGTCAGGCCCATCTGCTGCCCTTCTTTGATGAGGAGCTCACTCCAGCCACC 960
Qy 321 ProLeuTrpIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIle 340
Db 961 CCACCTCTGGATCATTTGGATGGGCTTACGAAGCAGAGTGGAGGGAAGATGCTGCACATA 1020
Qy 341 LeuLeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaTyr 360
Db 1021 CTGCTGCAGGCGTCAGTCAGGTCAATGACAGCACACGTCGCAATGCAGACGATCGGTAC 1080
Qy 361 GlnGlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThr 380
Db 1081 CAGGGGGAAGTCACCGAAGAAGATGATGTGCGAGGCATCCCGGAAGGGGGTGTGGACACC 1140
Qy 381 CysGlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGly 400
Db 1141 TGCCAGGGTGACAGTGGTGGGCCCTGATACCAATCTGACCACTGGCATGTGTGGGCG 1200
Qy 401 IleValSerTrpGlyTyrGlyCysGlyProSerThrProGlyValTyrThrLysVal 420
Db 1201 ATCGTTAGCTGGGGCTATGGCTGCGGGGCGCGAGCACCCAGGAGTATACACCAAGGTC 1260
Qy 421 SerAlaTyrLeuAsnTrpIleTyrAsnValTrpLys 432
Db 1261 TCAGCCTATCTCACTGGATCTACAATGTCTGGAAG 1296

RESULT 13
LOCUS CQ869994 2627 bp DNA linear PAT 13-SEP-2004
DEFINITION Sequence 415 from Patent WO2004074320.
ACCESSION CQ869994
VERSION CQ869994.1 GI:51999710

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1
AUTHORS Morris,D.W., Morris,D.W. and Malandro,M.S.
TITLE Novel therapeutic targets in cancer
JOURNAL Patent: WO 2004074320-A 415 02-SEP-2004;
Sagres Discovery, Inc. (US)

FEATURES
source
1..2627
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 1.88e-206 Length: 2627
Score: 2328.00 Matches: 433
Percent Similarity: 99.77% Conservative: 0
Best Local Similarity: 99.77% Mismatches: 1
Query Match: 99.40% Indels: 0
DB: 6 Gaps: 0

US-10-803-530-2 (1-435) x CQ869994 (1-2627)

Qy 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysProArg 21
Db 266 GATCTGCAGTGATCAACCTCTGACAGCCTCGATGTCAAAACCCCTGCGCAAAACCCGT 325
Qy 22 IleProMetGluThrPheArgLysValGlyIleProIleIleAlaLeuLeuSerLeu 41

Db 326 ATCCCCATGGAGACCTTCAGAAAGGTGGGATCCCCATCATCATAGCACTACTTGACCTG 385
Qy 42 AlaSerIleIleIleValValLeuIleLysValIleLeuAspLysTyrTyrPheLeu 61
Db 386 CGAGATATCATCATTTGGTTGTCTCATCAAGGTGATTCTGGATAAATACTACTTCTCTC 445
Qy 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGlyLeuAspCys 81
Db 446 TCGGGGAGCTCTCCACTTCATCCGAGGAAGCAGCTGTGTGACGGAGAGCTGACACTGT 505
Qy 82 ProLeuGlyGluAspGluHisCysValLysSerPheProGluGlyProAlaValAla 101
Db 506 CCTTGGGGGAGGAGGAGACACTGTGTCAAGAGCTTCCCGAAGGCGCTGCAGTGGCA 565
Qy 102 ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp 121
Db 566 GTCCGCTCTCCAAAGGACCGATCCACTGTCAGGTGTGGACTCGGCCACAGGAACCTGG 625
Qy 122 PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
Db 626 TTCTCTGCTGTTTCGACAACTTCACGAAGCTCTCGCTGAGACAGCCTGTAGGCAGATG 685
Qy 142 GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
Db 686 GGCTACAGCAGCAAAACCCACTTTCAGAGCTGTGAGATTGGCCAGACAGGATCTGGAT 745
Qy 162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
Db 746 GTTGTGTAATCAGCAAAACAGCAGGAGCTTCGATCGGAACTCAAGTGGGCGCTGT 805
Qy 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro 201
Db 806 CTCTCAGGCTCCCTGGTCTCCCTGCTGCTGCTTGGCTGTGGGAAGAGCTGAAGACCCCC 865
Qy 202 ArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
Db 866 CGTGTGGTGGGTGGAGGAGGCTCTGTGATCTTGGCTTGGCAGGTTCAGCATCCAG 925
Qy 222 TyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAla 241
Db 926 TACGACAAACAGCAGCTGTGGAGGAGCATCTGGACCCCTCCCTGCTTCACGGCA 985
Qy 242 AlahisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
Db 986 GCCCACTGCTTCAGGAAACATACCGATGTGTTCAACTGGAAGTGGGGGAGGCTCAGAC 1045
Qy 262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleIleGluPheAsnPro 281
Db 1046 AAACCTGGGCGAGCTTCCCATCCCTGCTGTGGCCAAAGATCATCATTTGAATTCACCCC 1105
Qy 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db 1106 ATGTACCCCAAGACAATGACATGCCCTCATGAAGCTGCAGTTCCTCCACTCACTTCTCA 1165
Qy 302 GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro 321
Db 1166 GGCACAGTCAGGCCCATCTGTCTGCCCTTCTTTGATGAGGAGCTCACTCCAGCCACCCA 1225
Qy 322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu 341
Db 1226 CTCTGGATCATTTGGATGGGCTTTACGAAGCAGAGTGGAGGGAAGATGTCTGACATACTG 1285
Qy 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaTyrGln 361
Db 1286 CTGCAGGCGTCAGTCCAGGTCAATTGACAGCACACCGTCAATGCGAGACGATCGTACCA 1345
Qy 362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys 381
Db 1346 GGGGAAGTCACCGAAGATGATGTGCGAGCATCTCCCGAAGGGGGTGTGGACACCTGC 1405
Qy 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle 401

Db	1406	CAGGGTGACAGTGGTGGGCCCTGATGATACCAATCTGACCAGTGGCATGTGGTGGGCATC	1465
Qy	402	ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSer	421
Db	1466	GTTAGTTGGGGCTATGGCTCGGGGGCCCGAGACCCAGGAGTATACCAAGGTCTCA	1525
Qy	422	AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu	435
Db	1526	GCCTATCTCACTGGATCTCAATGTTCTGGAAGGCTGAGCTG	1567
RESULT 14			
BD235831			
LOCUS			
DEFINITION	BD235831	2070 bp DNA linear	PAT 17-JUL-2003
ACCESSION		A novel method of diagnosing, monitoring, staging, imaging and	
VERSION	BD235831	treating various cancers.	
KEYWORDS	BD235831.1	GI:33045601	
SOURCE	JP 2002523760-A/2.		
ORGANISM	Homo sapiens (human)		
	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
	Hominidae; Homo.		
REFERENCE	1 (bases 1 to 2070)		
AUTHORS	Salceda,S., Sun,Y., Recipon,H. and Cafferkey,R.		
TITLE	A novel method of diagnosing, monitoring, staging, imaging and		
JOURNAL	treating various cancers		
	Patent: JP 2002523760-A 2 30-JUL-2002;		
COMMENT	DIADEXUS INC		
	OS Homo sapiens (human)		
	PN JP 2002523760-A/2		
	PD 30-JUL-2002		
	PF 01-SEP-1999 JP 2000567741		
	PI 02-SEP-1998 US 60/098880		
	PI SUSANA SALCEDA,YONGMING SUN,HERVE RECIPON,ROBERT CAFFERKEY PC		
	GOIN33/574,A61K39/395,A61K39/395,A61K39/395,A61K39/395 PC		
	,A61K49/00,A61K49/00,		
	PC A61K51/00,C07K16/32,C12N15/09,C12N15/00,A61K49/02 CC		
	A novel method of diagnosing, monitoring,		
	staging, imaging and		
	CC treating		
	CC various cancers		
FEATURES	PH Key	Location/Qualifiers	
FT	source	1..2070	
FT	source	/organism='Homo sapiens (human)'	
	Location/Qualifiers		
	1..2070		
	/organism='Homo sapiens'		
	/mol_type='genomic DNA'		
	/db_xref='taxon:9606'		
ORIGIN			
Alignment Scores:			
Pred. No.:	3.28e-206	Length:	2070
Score:	2324.00	Matches:	434
Percent Similarity:	99.77%	Conservative:	0
Best Local Similarity:	99.77%	Mismatches:	0
Query Match:	99.23%	Indels:	1
DB:	6	Gaps:	0
US-10-803-530-2 (1-435) x BD235831 (1-2070)			
Qy	2	AspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysProArg	21
Db	223	GATCCTGACAGTGATCAACCTCTGACAGAGCTCTGATGTCAAAACCCCTGGCAACCCCGT	282
Qy	22	IleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSerLeu	41
Db	283	ATCCCCATGGAGACCTTCAGAAAGGTGGGATCCCCATCATATAGCACTACTGACCTG	342
Qy	42	AlaSerIleIleIleValValValLysValLysValLysValLysValLysValLysVal	61
Db	343	CGGAGTATCATCTTGTTGGTTTGCTTCATCAGGTGATCTGGATAAATACTACTCTCTC	402
Qy	62	CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys	81
Db	403	TGCGGGCAGCCTCTCCACTTCATCCCGAGGAAGCAGCTGTGTGACGAGAGCTGGACTGT	462
Qy	82	ProLeuGlyGluAspGluHisCysValLysSerPheProGluGlyProAlaValAla	101
Db	463	CCCTTGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTCCCCGAGAGGCTTGCAGTGGCA	522
Qy	102	ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp	121
Db	523	GTCCGCCTCTCCAAAGGACCGATCCACACTGCAGGTCTGCAGCTCGGGCACAGGSACTGG	582
Qy	122	PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet	141
Db	583	TTCTCTGCTGTTTCGACAACTTCACAGAGCTCTCGCTGAGACAGCCTTAGCAGATG	642
Qy	142	GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp	161
Db	643	GGCTACAGCAGCAAAACCACTTTTCAGAGCTGTGAGATTGGCCACAGCAGGATCTGGAT	702
Qy	162	ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys	181
Db	703	GTGTGTTGAAATCAGAAAAACAGCCAGGAGCTTCGCATGCGGAACCTCAAGTGGGCCCTGT	762
Qy	182	LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro	201
Db	763	CTCTCAGGCTCCCTGGTCTCCCTGCACTGTCTTGCTGTGGGAAGAGCCTGAAGACCCC	822
Qy	202	ArgValValGlyGlyGluAlaSerValAspSerTrpProTrpGlnValSerIleGln	221
Db	823	CGTGTGTGGTGGGAGGAGGCTCTGTGGATTCTTTGGCTTTGGCAGTCCAGTCCAG	882
Qy	222	TyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThr-Al	241
Db	883	TACGACAAACAGCACCTCTGTGGAGGAGCATCTCTGGACCCCACTTGGGTCTCCACGGGC	942
Qy	241	aAlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAs	261
Db	943	AGCCCACTGCTTCAGGAACATACCATGTGTTCACCTGGAAGGTGGGGCAGGCTCAGA	1002
Qy	261	plysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsnPr	281
Db	1003	CAAACTGGGCAGCTTCCCTCCCTGCTGTGGCCCAAGATCATCATTTGAATTCAACCC	1062
Qy	281	oMetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSe	301
Db	1063	CATGTACCCCAAGACATGACATGCCCTCATGAAGCTGCAGTTCCTCCACTCCTTTCTC	1122
Qy	301	rGlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPr	321
Db	1123	AGGCACAGTCAGGCCCATCTGTCTGCCCTTCTTTGATGAGGAGCTCATCTCCAGCCACCCC	1182
Qy	321	oLeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLe	341
Db	1183	ACTTGGATCATTTGGATGGGCTTTTACGAAGCAGAAATGGAGGAGAGATGTCTGACATACT	1242
Qy	341	uLeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGl	361
Db	1243	GCTGACGGCTCAGTCCAGGTTCATTGACGACACAGCGTGCATTCAGACAGTGCATCA	1302
Qy	361	nglyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCy	381
Db	1303	GGGGGAAGTCAACGAGAAGATGATGTGTGCAGGCATCCCGGAAGGGGTGTGGACACCTG	1362
Qy	381	sglGlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIl	401
Db	1363	CCAGGGTGACAGTGTGGGGCCCCCTGATGATCAATCTGACAGTGGCATGTGGTGGGCAT	1422
Qy	401	eValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSe	421
Db	1423	CGTGTGGGCTATGGCTGCGGGGGCCCGAGACCCCGAGAGATATACACCAAGGTCTC	1482

```
QY 421 rAlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db 1483 AGCCTATCTCAACTGGATCTACAATGTCTGGAAGGCTGAGCTG 1525

RESULT 15
LOCUS AR232520 2079 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 1 from patent US 6455668.
ACCESSION AR232520
VERSION AR232520.1 GI:27274657
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 2079)
AUTHORS Mack,D., Gish,K.C. and Wilson,K.E.
TITLE Methods of diagnosing colorectal cancer, compositions, and methods
of screening for colorectal cancer modulators
JOURNAL Patent: US 6455668-A 1 24-SEP-2002;
Eos Biotechnology, Inc.; South San Francisco, CA
FEATURES
Location/Qualifiers
1..2079
/mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.: 3.3e-206 Length: 2079
Score: 2324.00 Matches: 434
Percent Similarity: 99.77% Conservative: 0
Best Local Similarity: 99.77% Mismatches: 0
Query Match: 99.23% Indels: 1
DB: 6 Gaps: 0

US-10-803-530-2 (1-435) x AR232520 (1-2079)

QY 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysProArg 21
Db 217 GATCTGACAGTGATCAACCTCTGACACAGCTCGAGTGTCAAAACCCCTGCGCAACCCCGT 276
QY 22 IleProMetGluThrPheArgLys-ValGlyIleProIleIleIleAlaLeuSerIle 41
Db 277 ATCCCATGGAGACCTTCAGAAAGTGTGGGATCCCATCATCATAGCACTACTGAGCCT 336
QY 41 uAlaSerIleIleIleValValLeuIleLysValIleLeuAspLysTyrTyrPheIle 61
Db 337 GCGAGATATCATATTGTGTTGCTCTCATCAAGTGATTCTGGATAAATACTACTTCTCT 396
QY 61 uCysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCy 81
Db 397 CTGCGGGCAGCCTCTCCACTTCATCCCGAGGAAGCAGCTGTGTGACGGAGAGCTGGACTG 456
QY 81 eProLeuGlyGluAspGluGluHisCysValLysSerPheProGluGlyProAlaValAl 101
Db 457 TCCCTTTGGGGAGAGCAGAGAGACATGTGTCAAGAGCTTCCCGAAGGGCCCTGCGAGTGGC 516
QY 101 aValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTr 121
Db 517 AGTCGCGCTCTCCAGGACCGATCCACATCGAGGTGTGGACTCGGCCACAGGAACCTG 576
QY 121 pPheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMe 141
Db 577 GTTCTCTGCTGTTTCGACAACTTCACAGAAGCTCTCGCTGAGACAGCCTGTAGGCAGAT 636
QY 141 tGlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAs 161
Db 637 GGGCTACACAGCAAAACCCACTTTTACAGACTGTGGAGATTGGCCACAGACCCAGGATCTGA 696
QY 161 pValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCy 181
Db 697 TGTGTTGGAATACAGAAAAACAGCCAGGAGCTTGCAATGCCGAACCTCAAGTGGGCCCTG 756
QY 181 sLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPr 201
```

```
Db 757 TCTCTCAGGCTCCCTGGTCTCCCTGCACTGTCTGCTCTGGGAAGACCTGAAGACCCC 816
QY 201 oArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleGl 221
Db 817 CCGTGTGTGGTGGGAGGAGGCTCTGTGGATTCTTGGCTTGGCAGGTGAGCATCCA 876
QY 221 nTyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAl 241
Db 877 GTACGACAAACAGCACGCTCTGTGAGGGAGCATCTCTGACCCCACTGGGTCTCACGGC 936
QY 241 aAlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValAAlaGlySerAs 261
Db 937 AGCCCATCTGCTTCAGAAACATACCGATGTGTCAACTGGAAGGTGCGGAGGCTCAGA 996
QY 261 pLysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleIleGluPheAsnPr 281
Db 997 CAAACTGGGAGCTTCCCATCCCTGGCTGTGGCCAAAGATCATCATTAATTAATCAAGCC 1056
QY 281 oMetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSe 301
Db 1057 CATGTACCCCAAGACAAATGACATCGCCCTCATGAAGCTGCAGTTCCCACTCACTTCTC 1116
QY 301 rGlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPr 321
Db 1117 AGGCACAGTCAGGCCCATCTGTCTGCCCTTTCTTGATGAGGAGCTCATCTCCAGCCACCCC 1176
QY 321 oLeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLe 341
Db 1177 ACTCTGGATCATTTGATGGGCTTTTACGAAGCAGAAATGGAGGAAGATGTCTGACATACT 1236
QY 341 uLeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaTyrGl 361
Db 1237 GCTGCAGGCGTCAGTCCAGGTCAATTGACAGCACACGGTGCATGCGAGATGCGTACCA 1296
QY 361 nGlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCy 381
Db 1297 GGGGAGTCAACCGAAGATGATGTGTGAGGCAATCCCGAGGGGGGTGTGGACACCTG 1356
QY 381 sGlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyI 401
Db 1357 CCAGSGTGACAGTGGTGGGCCCTGATGTACCAATCTGACCAGTGGCATGTGGTGGGCAT 1416
QY 401 eValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSe 421
Db 1417 CGTTAGCTGGGCTATGGCTGCGGGGCCCGAGCACCCCGAGAGTATACACCAAGGTCTC 1476
QY 421 rAlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db 1477 AGCCTATCTCAACTGGATCTACAATGTCTGGAAGGCTGAGCTG 1519
```

Search completed: December 6, 2005, 21:56:59

Job time : 5677 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2005, 13:53:15 ; Search time 166 Seconds
(without alignments)
1094.914 Million cell updates/sec

Title: US-10-803-530-2
Perfect score: 2342
Sequence: 1 MDPDSQPLNSLDVKPLRKP.....VTKVSALNWIYVWKAEL 435

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA Main:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2342	100.0	435	US-10-803-530-2	Sequence 2, Appli
2	2342	100.0	435	US-10-473-127-1578	Sequence 1578, Ap
3	2342	100.0	435	US-10-473-127-1597	Sequence 1597, Ap
4	2338	99.8	435	US-09-968-415-6	Sequence 6, Appli
5	2338	99.8	435	US-10-180-719-6	Sequence 6, Appli
6	2338	99.8	435	US-10-473-127-1596	Sequence 1596, Ap
7	2338	99.8	435	US-11-045-577-6	Sequence 6, Appli
8	2337	99.8	437	US-09-851-588-8	Sequence 8, Appli
9	2337	99.8	437	US-10-295-027-779	Sequence 779, App
10	2337	99.8	437	US-10-295-027-791	Sequence 791, App
11	2337	99.8	437	US-10-295-027-831	Sequence 831, App
12	2337	99.8	437	US-10-295-027-1196	Sequence 1196, Ap
13	2337	99.8	437	US-10-173-999-89	Sequence 89, Appli
14	2337	99.8	437	US-10-473-127-1581	Sequence 1581, Ap
15	2337	99.8	437	US-10-473-127-1586	Sequence 1586, Ap
16	2337	99.8	437	US-10-473-127-1601	Sequence 1601, Ap
17	2337	99.8	437	US-10-473-127-1602	Sequence 1602, Ap
18	2337	99.8	437	US-10-994-117-7	Sequence 7, Appli
19	2337	99.8	437	US-10-991-287-7	Sequence 7, Appli
20	2329	99.4	492	US-10-473-127-1579	Sequence 1579, Ap
21	2329	99.4	492	US-10-473-127-1598	Sequence 1598, Ap
22	2319	99.0	437	US-09-776-191-4	Sequence 4, Appli
23	2319	99.0	437	US-10-156-214A-4	Sequence 4, Appli
24	2319	99.0	437	US-10-473-127-1583	Sequence 1583, Ap
25	2319	99.0	437	US-10-473-127-1588	Sequence 1588, Ap
26	2297.5	98.1	432	US-09-888-257A-7	Sequence 7, Appli
27	2297.5	98.1	432	US-09-946-374-275	Sequence 275, App

28	2297.5	98.1	432	4	US-10-006-867-112	Sequence 112, App
29	2297.5	98.1	432	4	US-10-052-586-330	Sequence 330, App
30	2297.5	98.1	432	4	US-10-063-547-112	Sequence 112, App
31	2297.5	98.1	432	4	US-10-063-551-112	Sequence 112, App
32	2297.5	98.1	432	4	US-10-174-590-330	Sequence 330, App
33	2297.5	98.1	432	4	US-10-176-758-330	Sequence 330, App
34	2297.5	98.1	432	4	US-10-175-737-330	Sequence 330, App
35	2297.5	98.1	432	4	US-10-063-616-112	Sequence 112, App
36	2297.5	98.1	432	4	US-10-174-581-330	Sequence 330, App
37	2297.5	98.1	432	4	US-10-176-483-330	Sequence 330, App
38	2297.5	98.1	432	4	US-10-176-749-330	Sequence 330, App
39	2297.5	98.1	432	4	US-10-176-914-330	Sequence 330, App
40	2297.5	98.1	432	4	US-10-176-915-330	Sequence 330, App
41	2297.5	98.1	432	4	US-10-063-569-112	Sequence 112, App
42	2297.5	98.1	432	4	US-10-063-513-112	Sequence 112, App
43	2297.5	98.1	432	4	US-10-063-515-112	Sequence 112, App
44	2297.5	98.1	432	4	US-10-063-512-112	Sequence 112, App
45	2297.5	98.1	432	4	US-10-173-706-330	Sequence 330, App

ALIGNMENTS

RESULT 1

US-10-803-530-2
; Sequence 2, Application US/10803530
; Publication No. US20040146981A1
; GENERAL INFORMATION:
; APPLICANT: Darrow, Andrew L
; APPLICANT: Qi, Jain-shen
; APPLICANT: Andrade-Gordon, Patricia
; TITLE OF INVENTION: DNA encoding human serine protease D-G
; FILE REFERENCE: ORT-1273
; CURRENT APPLICATION NUMBER: US/10/803,530
; CURRENT FILING DATE: 2004-03-17
; PRIOR APPLICATION NUMBER: US/09/607,745
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-803-530-2

Query Match	100.0%;	Score 2342;	DB 4;	Length 435;
Best Local Similarity	100.0%;	Pred. No. 1.9e-219;	Mismatches 0;	Indels 0;
Matches 435;	Conservative 0;			Gaps 0;
Qy	1	MDPDSQPLNSLDVKPLRKPIMETFRKVGIPITIIALLSLASIIIVVVLKIVLDKYYF	60	
Db	1	MDPDSQPLNSLDVKPLRKPIMETFRKVGIPITIIALLSLASIIIVVVLKIVLDKYYF	60	
Qy	61	LCGQPLHPIPRKQDCGELDCPLGDEBHCVKSPFEGPAVRLSKDRSTLOVLSATGN	120	
Db	61	LCGQPLHPIPRKQDCGELDCPLGDEBHCVKSPFEGPAVRLSKDRSTLOVLSATGN	120	
Qy	121	WFSACDFNTEALATACRMGYSSKPTFRAVEIGPDQDLDVVEITENSQELMRNSSGP	180	
Db	121	WFSACDFNTEALATACRMGYSSKPTFRAVEIGPDQDLDVVEITENSQELMRNSSGP	180	
Qy	181	CLSGSLVSLHCLACGKSLKTPRVVGEEASVDSWPQVSIQYDKQHVCGGSLDPHWLVT	240	
Db	181	CLSGSLVSLHCLACGKSLKTPRVVGEEASVDSWPQVSIQYDKQHVCGGSLDPHWLVT	240	
Qy	241	AAHCFRKHTDVFNWVKVRAGSKLGSPSLAVAKIIIEFNPMYKXNDITALKLOPPLTF	300	
Db	241	AAHCFRKHTDVFNWVKVRAGSKLGSPSLAVAKIIIEFNPMYKXNDITALKLOPPLTF	300	
Qy	301	SGTVRPICLPFFDEELTPATPLWIIWGFTKONGKMSDILLOASVQVIDSTRCNADDAY	360	
Db	301	SGTVRPICLPFFDEELTPATPLWIIWGFTKONGKMSDILLOASVQVIDSTRCNADDAY	360	

QY 361 QGEVTERKMCAGIPEGGVDTCCGDSGGPLMYQSDQWHVVGIVSWGCGGSPSTPGVYTKV 420
Db |||||||
361 QGEVTERKMCAGIPEGGVDTCCGDSGGPLMYQSDQWHVVGIVSWGCGGSPSTPGVYTKV 420
QY 421 SAYLNWIYVWKAE 435
Db |||||||
421 SAYLNWIYVWKAE 435

RESULT 2
US-10-473-127-1578
; Sequence 1578, Application US/10473127
; Publication No. US20040236091A1
; GENERAL INFORMATION:
; APPLICANT: Zycos Inc.
; TITLE OF INVENTION: TRANSLATIONAL PROFILING
; FILE REFERENCE: 08191-026W01
; CURRENT APPLICATION NUMBER: US/10/473,127
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 60/279,495
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: 60/292,544
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 60/310,801
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/326,370
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: 60/336,780
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/358,985
; PRIOR FILING DATE: 2002-02-20
; NUMBER OF SEQ ID NOS: 2041
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1578
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-473-127-1578

Query Match 100.0%; Score 2342; DB 5; Length 435;
Best Local Similarity 100.0%; Pred. No. 1.9e-219;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDPDSQPLNSLDVVKPLRPKIPMETFRKVGIPPIIALLSLASIIIVVLIKVILDKYF 60
Db |||||||
1 MDPDSQPLNSLDVVKPLRPKIPMETFRKVGIPPIIALLSLASIIIVVLIKVILDKYF 60
QY 61 LCGQPLHPIPRKQLCDGELDCPLGEDEHCVKSPFEGPAPAVRLSKDRSTLQVLD SATGN 120
Db |||||||
61 LCGQPLHPIPRKQLCDGELDCPLGEDEHCVKSPFEGPAPAVRLSKDRSTLQVLD SATGN 120
QY 121 WFSACFDNFTALAEACROMGYSSKPTFRAVEIGPDQDLVVEITENSQELMRNSSGP 180
Db |||||||
121 WFSACFDNFTALAEACROMGYSSKPTFRAVEIGPDQDLVVEITENSQELMRNSSGP 180
QY 181 CLSGSLVSLHCLACGKSLKTPRVVGGEEASVDSWPQVSIQYDKQHVCGGSILDPHWLT 240
Db |||||||
181 CLSGSLVSLHCLACGKSLKTPRVVGGEEASVDSWPQVSIQYDKQHVCGGSILDPHWLT 240
QY 241 AAHCFRKHDTVFNKVRAGSKLGSFPFLSAVAKIIIEFNPMYPKNDIALMKLQPLTF 300
Db |||||||
241 AAHCFRKHDTVFNKVRAGSKLGSFPFLSAVAKIIIEFNPMYPKNDIALMKLQPLTF 300
QY 301 SGTVRPICLPFFDEBELTPATPLWIIWGFTKQNGKMSDILLQASVQVIDSTRCNADDA 360
Db |||||||
301 SGTVRPICLPFFDEBELTPATPLWIIWGFTKQNGKMSDILLQASVQVIDSTRCNADDA 360
QY 361 QGEVTERKMCAGIPEGGVDTCCGDSGGPLMYQSDQWHVVGIVSWGCGGSPSTPGVYTKV 420
Db |||||||
361 QGEVTERKMCAGIPEGGVDTCCGDSGGPLMYQSDQWHVVGIVSWGCGGSPSTPGVYTKV 420
QY 421 SAYLNWIYVWKAE 435
Db |||||||
421 SAYLNWIYVWKAE 435

RESULT 4
US-09-968-415-6
; Sequence 6, Application US/09968415

Db 421 SAYLNWIYVWKAE 435
RESULT 3
US-10-473-127-1597
; Sequence 1597, Application US/10473127
; Publication No. US20040236091A1
; GENERAL INFORMATION:
; APPLICANT: Zycos Inc.
; TITLE OF INVENTION: TRANSLATIONAL PROFILING
; FILE REFERENCE: 08191-026W01
; CURRENT APPLICATION NUMBER: US/10/473,127
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 60/279,495
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: 60/292,544
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 60/310,801
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/326,370
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: 60/336,780
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/358,985
; PRIOR FILING DATE: 2002-02-20
; NUMBER OF SEQ ID NOS: 2041
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1597
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-473-127-1597

Query Match 100.0%; Score 2342; DB 5; Length 435;
Best Local Similarity 100.0%; Pred. No. 1.9e-219;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDPDSQPLNSLDVVKPLRPKIPMETFRKVGIPPIIALLSLASIIIVVLIKVILDKYF 60
Db |||||||
1 MDPDSQPLNSLDVVKPLRPKIPMETFRKVGIPPIIALLSLASIIIVVLIKVILDKYF 60
QY 61 LCGQPLHPIPRKQLCDGELDCPLGEDEHCVKSPFEGPAPAVRLSKDRSTLQVLD SATGN 120
Db |||||||
61 LCGQPLHPIPRKQLCDGELDCPLGEDEHCVKSPFEGPAPAVRLSKDRSTLQVLD SATGN 120
QY 121 WFSACFDNFTALAEACROMGYSSKPTFRAVEIGPDQDLVVEITENSQELMRNSSGP 180
Db |||||||
121 WFSACFDNFTALAEACROMGYSSKPTFRAVEIGPDQDLVVEITENSQELMRNSSGP 180
QY 181 CLSGSLVSLHCLACGKSLKTPRVVGGEEASVDSWPQVSIQYDKQHVCGGSILDPHWLT 240
Db |||||||
181 CLSGSLVSLHCLACGKSLKTPRVVGGEEASVDSWPQVSIQYDKQHVCGGSILDPHWLT 240
QY 241 AAHCFRKHDTVFNKVRAGSKLGSFPFLSAVAKIIIEFNPMYPKNDIALMKLQPLTF 300
Db |||||||
241 AAHCFRKHDTVFNKVRAGSKLGSFPFLSAVAKIIIEFNPMYPKNDIALMKLQPLTF 300
QY 301 SGTVRPICLPFFDEBELTPATPLWIIWGFTKQNGKMSDILLQASVQVIDSTRCNADDA 360
Db |||||||
301 SGTVRPICLPFFDEBELTPATPLWIIWGFTKQNGKMSDILLQASVQVIDSTRCNADDA 360
QY 361 QGEVTERKMCAGIPEGGVDTCCGDSGGPLMYQSDQWHVVGIVSWGCGGSPSTPGVYTKV 420
Db |||||||
361 QGEVTERKMCAGIPEGGVDTCCGDSGGPLMYQSDQWHVVGIVSWGCGGSPSTPGVYTKV 420
QY 421 SAYLNWIYVWKAE 435
Db |||||||
421 SAYLNWIYVWKAE 435

Db	301	SGVVRPCLPFDFBELPATPLWIIWGFTKQNGKMSDILLQASVQVIDSTRCNADDAY	360
Qy	361	QGEYTERKMCAGIPEGGVDTCCQDSGGPLMTYQSDQWVHVGVISWGYGCGPSTPGVYTKV	420
Db	361	QGEYTERKMCAGIPEGGVDTCCQDSGGPLMTYQSDQWVHVGVISWGYGCGPSTPGVYTKV	420
Qy	421	SAYLNWYINVKAEKEL 435	
Db	421	SAYLNWYINVKAEKEL 435	
RESULT 5			
US-10-180-719-6			
; Sequence 6, Application US/10180719			
; Publication No. US20030166246A1			
; GENERAL INFORMATION:			
; APPLICANT: Bandman, Olga			
; Hillman, Jennifer L.			
; Yue, Henry			
; Guegler, Karl J.			
; Corley, Neil C.			
; Tang, Tom Y.			
; Shah, Purvi			
; TITLE OF INVENTION: HUMAN PROTEASE MOLECULES			
; NUMBER OF SEQUENCES: 24			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Incyte Pharmaceuticals, Inc.			
; STREET: 3174 Porter Dr.			
; CITY: Palo Alto			
; STATE: CA			
; COUNTRY: USA			
; ZIP: 94304			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Diskette			
; COMPUTER: IBM Compatible			
; OPERATING SYSTEM: DOS			
; SOFTWARE: FastSeq for Windows Version 2.0			
; CURRENT APPLICATION DATA:			
; FILING APPLICATION NUMBER: US/10/180,719			
; FILING DATE: 25-Jun-2002			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: US/09/008,271			
; FILING DATE: 16-Jan-1998			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Mohan-Peterson, Sheela			
; REGISTRATION NUMBER: 41,201			
; REFERENCE/DOCKET NUMBER: PF-0458 US			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: 650-855-0555			
; TELEFAX: 650-845-4166			
; INFORMATION FOR SEQ ID NO: 6:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 435 amino acids			
; TYPE: amino acid			
; STRANDEDNESS: single			
; TOPOLOGY: linear			
; IMMEDIATE SOURCE:			
; LIBRARY: COLNNOT13			
; CLONE: 1337018			
; SEQUENCE DESCRIPTION: SEQ ID NO: 6 :			
US-10-180-719-6			
Query Match 99.8%; Score 2338; DB 4; Length 435;			
Best Local Similarity 99.8%; Pred. No. 4.7e-219;			
Matches 434; Conservative 1; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	MDPDSQPLNSLDVKPLKPRIPMETFRKVGPIIIALLSLASIIIVVLLIKVILDKYIF	60
Db	1	MDPDSQPLNSLDVKPLKPRIPMETFRKVGPIIIALLSLASIIIVVLLIKVILDKYIF	60
Qy	61	LCGQPLHPIPRKQLCDGELDCPLGDEDEHCVKSPFEGPAVAVRLSKORSTLQVLDSATGN	120

Db 61 LCGQLHFIPIRKQKQDGLDCLPGBDEEHCVKSPFEGPAVAVRLSKDRSTLQVLDSATGN 120
 Qy 121 WFSACDFNTEALAEACROMGYSSKPTFRAVEIGPDQDLDVVEITENSQELMRNSSGP 180
 Db 121 WFSACDFNTEALAEACROMGYSSKPTFRAVEIGPDQDLDVVEITENSQELMRNSSGP 180
 Qy 181 CLSGSLVSLHCLACGSKLTPRVVGGEEASVDSWPMQVSIQYDKQHVCGGSIIDPHHWLT 240
 Db 181 CLSGSLVSLHCLACGSKLTPRVVGGEEASVDSWPMQVSIQYDKQHVCGGSIIDPHHWLT 240
 Qy 241 AAHCFRKHITDVFNKVRAGSKLGSFPSPSLAVAKIIIEFNPMYPKNDIALMKLOFPLTF 300
 Db 241 AAHCFRKHITDVFNKVRAGSKLGSFPSPSLAVAKIIIEFNPMYPKNDIALMKLOFPLTF 300
 Qy 301 SGTVPICLPFDEBELTATPLWIIIGWFTKQNGKMSDILLQASVQVIDSTRCNADAY 360
 Db 301 SGTVPICLPFDEBELTATPLWIIIGWFTKQNGKMSDILLQASVQVIDSTRCNADAY 360
 Qy 361 QGEVTEKMWKAGIPGGVDTCQDGGGGLMYQSDQHWVGVISWGYCGGSPSTGVYTKV 420
 Db 361 QGEVTEKMWKAGIPGGVDTCQDGGGGLMYQSDQHWVGVISWGYCGGSPSTGVYTKV 420
 Qy 421 SAYLNWIYNVWKAEL 435
 Db 421 SAYLNWIYNVWKAEL 435

RESULT 6
 US-10-473-127-1596
 ; Sequence 1596, Application US/10473127
 ; Publication No. US20040236091A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Zycos Inc.
 ; TITLE OF INVENTION: TRANSLATIONAL PROFILING
 ; FILE REFERENCE: 08191-026W01
 ; CURRENT APPLICATION NUMBER: US/10/473,127
 ; CURRENT FILING DATE: 2003-09-26
 ; PRIOR APPLICATION NUMBER: 60/279,495
 ; PRIOR FILING DATE: 2001-03-28
 ; PRIOR APPLICATION NUMBER: 60/292,544
 ; PRIOR FILING DATE: 2001-05-21
 ; PRIOR APPLICATION NUMBER: 60/310,801
 ; PRIOR FILING DATE: 2001-08-08
 ; PRIOR APPLICATION NUMBER: 60/326,370
 ; PRIOR FILING DATE: 2001-10-01
 ; PRIOR APPLICATION NUMBER: 60/336,780
 ; PRIOR FILING DATE: 2001-12-04
 ; PRIOR APPLICATION NUMBER: 60/358,985
 ; PRIOR FILING DATE: 2002-02-20
 ; NUMBER OF SEQ ID NOS: 2041
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1596
 ; LENGTH: 435
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-473-127-1596

Query Match 99.8%; Score 2338; DB 5; Length 435;
 Best Local Similarity 99.8%; Pred. No. 4.7e-219;
 Matches 434; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MDPDSQPLNSLDVVKLPKRPIMETFRKVGIPITIIALLSLASIIIVVLIKVLIDKYF 60
 Db 1 MDPDSQPLNSLDVVKLPKRPIMETFRKVGIPITIIALLSLASIIIVVLIKVLIDKYF 60
 Qy 61 LCGQLHFIPIRKQKQDGLDCLPGBDEEHCVKSPFEGPAVAVRLSKDRSTLQVLDSATGN 120
 Db 61 LCGQLHFIPIRKQKQDGLDCLPGBDEEHCVKSPFEGPAVAVRLSKDRSTLQVLDSATGN 120
 Qy 121 WFSACDFNTEALAEACROMGYSSKPTFRAVEIGPDQDLDVVEITENSQELMRNSSGP 180
 Db 121 WFSACDFNTEALAEACROMGYSSKPTFRAVEIGPDQDLDVVEITENSQELMRNSSGP 180

Qy 181 CLSGSLVSLHCLACGSKLTPRVVGGEEASVDSWPMQVSIQYDKQHVCGGSIIDPHHWLT 240
 Db 181 CLSGSLVSLHCLACGSKLTPRVVGGEEASVDSWPMQVSIQYDKQHVCGGSIIDPHHWLT 240
 Qy 241 AAHCFRKHITDVFNKVRAGSKLGSFPSPSLAVAKIIIEFNPMYPKNDIALMKLOFPLTF 300
 Db 241 AAHCFRKHITDVFNKVRAGSKLGSFPSPSLAVAKIIIEFNPMYPKNDIALMKLOFPLTF 300
 Qy 301 SGTVPICLPFDEBELTATPLWIIIGWFTKQNGKMSDILLQASVQVIDSTRCNADAY 360
 Db 301 SGTVPICLPFDEBELTATPLWIIIGWFTKQNGKMSDILLQASVQVIDSTRCNADAY 360
 Qy 361 QGEVTEKMWKAGIPGGVDTCQDGGGGLMYQSDQHWVGVISWGYCGGSPSTGVYTKV 420
 Db 361 QGEVTEKMWKAGIPGGVDTCQDGGGGLMYQSDQHWVGVISWGYCGGSPSTGVYTKV 420
 Qy 421 SAYLNWIYNVWKAEL 435
 Db 421 SAYLNWIYNVWKAEL 435

RESULT 7
 US-11-045-577-6
 ; Sequence 6, Application US/11045577
 ; Publication No. US20050181404A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bandman, Olga
 ; Hillman, Jennifer L.
 ; Yue, Henry
 ; Guegler, Karl J.
 ; Corley, Neil C.
 ; Tang, Tom Y.
 ; Shah, Purvi
 ; TITLE OF INVENTION: HUMAN PROTEASE MOLECULES
 ; NUMBER OF SEQUENCES: 24
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Dr.
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/11/045,577
 ; FILING DATE: 27-Jan-2005
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/659,151
 ; FILING DATE: 11-Sep-2000
 ; APPLICATION NUMBER: 09/008,271
 ; FILING DATE: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Mohan-Peterson, Sheela
 ; REGISTRATION NUMBER: 41,201
 ; REFERENCE/DOCKET NUMBER: PF-0458 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 650-855-0555
 ; TELEFAX: 650-845-4166
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 435 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: COLNNOT13
 ; CLONE: 1337018
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 6 :
 US-11-045-577-6

```
Query Match          99.8%; Score 2338; DB 6; Length 435;
Best Local Similarity 99.8%; Pred. No. 4.7e-219;
Matches 434; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPDSQPLNSLDVPLKRPIMETFRKVGIPPIIALLSLASIIIVVLKIVLDKYIF 60
DB 1 MDPDSQPLNSLDVPLKRPIMETFRKVGIPPIIALLSLASIIIVVLKIVLDKYIF 60

QY 61 LCGQPLHFI PRKQLCDGELDCPLGDEBHCVKSPFEGPAVAVRLSKDRSTLQVLD SATGN 120
DB 61 LCGQPLHFI PRKQLCDGELDCPLGDEBHCVKSPFEGPAVAVRLSKDRSTLQVLD SATGN 120

QY 121 WFSACDFNFTALAEATACRQMGYSKPTFRAVEIGPDQDLVDVEITENSQELMRNSSGP 180
DB 121 WFSACDFNFTALAEATACRQMGYSKPTFRAVEIGPDQDLVDVEITENSQELMRNSSGP 180

QY 181 CLSGSLVSLHCLACGSKSLKTPRVVGGEASVDSWPQVSIQYDKQHVCGGSI LDPHWWLT 240
DB 181 CLSGSLVSLHCLACGSKSLKTPRVVGGEASVDSWPQVSIQYDKQHVCGGSI LDPHWWLT 240

QY 241 AAHCFRKHDTVFNWVKRAGSKLGFSPSLAVAKIIIEFNPMYPKNDIALMKLQFP LTF 300
DB 241 AAHCFRKHDTVFNWVKRAGSKLGFSPSLAVAKIIIEFNPMYPKNDIALMKLQFP LTF 300

QY 301 SGTVPICLPFDEBLTATPLWIIWGFTKONGKMSDILLOASVQVIDSTRCNAD DAY 360
DB 301 SGTVPICLPFDEBLTATPLWIIWGFTKONGKMSDILLOASVQVIDSTRCNAD DAY 360

QY 361 QGEVTEKMWKACGIPGGVDTCQDGGPLMYQSDQHVVGIVSWGCGGPGSTPGVYTKV 420
DB 361 QGEVTEKMWKACGIPGGVDTCQDGGPLMYQSDQHVVGIVSWGCGGPGSTPGVYTKV 420

QY 421 SAYLWNIYNVWKAEL 435
DB 421 SAYLWNIYNVWKAEL 435

RESULT 8
US-09-851-588-8
; Sequence 8, Application US/09851588
; Patent No. US20020042067A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David
; APPLICANT: Gish, Kurt C.
; APPLICANT: Wilson, Keith E.
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSING COLORECTAL CANCER, COMPOSITIONS, AND
; FILE REFERENCE: A-68829-1/DJB/JJD/AMS
; CURRENT APPLICATION NUMBER: US/09/851,588
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 09/642,252
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: US 09/656,002
; PRIOR FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-851-588-8

Query Match          99.8%; Score 2337; DB 3; Length 437;
Best Local Similarity 100.0%; Pred. No. 5.9e-219;
Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DPDSQPLNSLDVPLKRPIMETFRKVGIPPIIALLSLASIIIVVLKIVLDKYIFL 61
DB 4 DPDSQPLNSLDVPLKRPIMETFRKVGIPPIIALLSLASIIIVVLKIVLDKYIFL 63

QY 62 CQGPLHFI PRKQLCDGELDCPLGDEBHCVKSPFEGPAVAVRLSKDRSTLQVLD SATGNW 121
```

```
DB 64 CQGPLHFI PRKQLCDGELDCPLGDEBHCVKSPFEGPAVAVRLSKDRSTLQVLD SATGNW 123
QY 122 FSACDFNFTALAEATACRQMGYSKPTFRAVEIGPDQDLVDVEITENSQELMRNSSGP 181
DB 124 FSACDFNFTALAEATACRQMGYSKPTFRAVEIGPDQDLVDVEITENSQELMRNSSGP 183
QY 182 LSGSLVSLHCLACGSKSLKTPRVVGGEASVDSWPQVSIQYDKQHVCGGSI LDPHWWLT 241
DB 184 LSGSLVSLHCLACGSKSLKTPRVVGGEASVDSWPQVSIQYDKQHVCGGSI LDPHWWLT 243
QY 242 AHCFRKHDTVFNWVKRAGSKLGFSPSLAVAKIIIEFNPMYPKNDIALMKLQFP LTF 301
DB 244 AHCFRKHDTVFNWVKRAGSKLGFSPSLAVAKIIIEFNPMYPKNDIALMKLQFP LTF 303
QY 302 GTVPICLPFDEBLTATPLWIIWGFTKONGKMSDILLOASVQVIDSTRCNAD DAY 361
DB 304 GTVPICLPFDEBLTATPLWIIWGFTKONGKMSDILLOASVQVIDSTRCNAD DAY 363
QY 362 GEVTEKMWKACGIPGGVDTCQDGGPLMYQSDQHVVGIVSWGCGGPGSTPGVYTKV 421
DB 364 GEVTEKMWKACGIPGGVDTCQDGGPLMYQSDQHVVGIVSWGCGGPGSTPGVYTKV 423
QY 422 AYLWNIYNVWKAEL 435
DB 424 AYLWNIYNVWKAEL 437

RESULT 9
US-10-295-027-779
; Sequence 779, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 779
; LENGTH: 437
; TYPE: PRT
```


; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 831
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-831

Query Match 99.8%; Score 2337; DB 4; Length 437;
Best Local Similarity 100.0%; Pred. No. 5.9e-219;
Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DPDSQPLNSLDVKEPRIPMETFRKVGIPPIIIALLSLASIIIVVILKILDKYFL 61
DB 4 DPDSQPLNSLDVKEPRIPMETFRKVGIPPIIIALLSLASIIIVVILKILDKYFL 63
QY 62 CGQPLHFTPRKQDCGELDCPLGEDEEHCHVKSPFEGPAVAVRLSKDRSTLQVLDATGNW 121
DB 64 CGQPLHFTPRKQDCGELDCPLGEDEEHCHVKSPFEGPAVAVRLSKDRSTLQVLDATGNW 123
QY 122 FSACFDNTEALAEACRQMGYSKPTFRAVEIGDPDQLDVVEITENSQELMRNSSGPC 181
DB 124 FSACFDNTEALAEACRQMGYSKPTFRAVEIGDPDQLDVVEITENSQELMRNSSGPC 183
QY 182 LSGSLVSLHCLACGSKSLKTPRVVGEEASVDSWPQVSIQYDKQHVCGGSIIDPHWVLT 241
DB 184 LSGSLVSLHCLACGSKSLKTPRVVGEEASVDSWPQVSIQYDKQHVCGGSIIDPHWVLT 243
QY 242 AHCPRKHTDVFNNKVRAGSKLGSFPLSAVAKIIIEFNPMYPKNDIALMKLQPLTF 301
DB 244 AHCPRKHTDVFNNKVRAGSKLGSFPLSAVAKIIIEFNPMYPKNDIALMKLQPLTF 303
QY 302 GTVRPCLPFFDEELTPATPLMIIGWFTKQNGKMSDILLQASVQVIDSTRCNADDA 361
DB 304 GTVRPCLPFFDEELTPATPLMIIGWFTKQNGKMSDILLQASVQVIDSTRCNADDA 363
QY 362 GEVTEKMMKAGIPEGGVDTCCQDSGGPLMYQSDQHVHVGIVSWGCGGSPSTPGVYTK 421
DB 364 GEVTEKMMKAGIPEGGVDTCCQDSGGPLMYQSDQHVHVGIVSWGCGGSPSTPGVYTK 423
QY 422 AYLNMIYNWKAEL 435
DB 424 AYLNMIYNWKAEL 437

RESULT 12

US-10-295-027-1196
; Sequence 1196, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natacha

; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1196
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-1196

Query Match 99.8%; Score 2337; DB 4; Length 437;
Best Local Similarity 100.0%; Pred. No. 5.9e-219;
Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DPDSQPLNSLDVKEPRIPMETFRKVGIPPIIIALLSLASIIIVVILKILDKYFL 61
DB 4 DPDSQPLNSLDVKEPRIPMETFRKVGIPPIIIALLSLASIIIVVILKILDKYFL 63
QY 62 CGQPLHFTPRKQDCGELDCPLGEDEEHCHVKSPFEGPAVAVRLSKDRSTLQVLDATGNW 121
DB 64 CGQPLHFTPRKQDCGELDCPLGEDEEHCHVKSPFEGPAVAVRLSKDRSTLQVLDATGNW 123
QY 122 FSACFDNTEALAEACRQMGYSKPTFRAVEIGDPDQLDVVEITENSQELMRNSSGPC 181
DB 124 FSACFDNTEALAEACRQMGYSKPTFRAVEIGDPDQLDVVEITENSQELMRNSSGPC 183
QY 182 LSGSLVSLHCLACGSKSLKTPRVVGEEASVDSWPQVSIQYDKQHVCGGSIIDPHWVLT 241
DB 184 LSGSLVSLHCLACGSKSLKTPRVVGEEASVDSWPQVSIQYDKQHVCGGSIIDPHWVLT 243
QY 242 AHCPRKHTDVFNNKVRAGSKLGSFPLSAVAKIIIEFNPMYPKNDIALMKLQPLTF 301
DB 244 AHCPRKHTDVFNNKVRAGSKLGSFPLSAVAKIIIEFNPMYPKNDIALMKLQPLTF 303
QY 302 GTVRPCLPFFDEELTPATPLMIIGWFTKQNGKMSDILLQASVQVIDSTRCNADDA 361
DB 304 GTVRPCLPFFDEELTPATPLMIIGWFTKQNGKMSDILLQASVQVIDSTRCNADDA 363
QY 362 GEVTEKMMKAGIPEGGVDTCCQDSGGPLMYQSDQHVHVGIVSWGCGGSPSTPGVYTK 421
DB 364 GEVTEKMMKAGIPEGGVDTCCQDSGGPLMYQSDQHVHVGIVSWGCGGSPSTPGVYTK 423

Qy 422 AYLNWIYNVWKAEL 435
Db 424 AYLNWIYNVWKAEL 437

RESULT 13

US-10-173-999-89
; Sequence 89, Application US/10173999
; Publication No. US20040005563A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Ovarian Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Ovarian
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002420US
; CURRENT APPLICATION NUMBER: US/10/173,999
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: US 60/299,234
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/315,287
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 163
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 89
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-173-999-89

Query Match 99.8%; Score 2337; DB 4; Length 437;
Best Local Similarity 100.0%; Pred. No. 5.9e-219;
Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 DPDSQPLNSLDVKLPKPRIPMETFRKVGIPPIIALLSLASIIIVVLIKVILDKYYFL 61
Db 4 DPDSQPLNSLDVKLPKPRIPMETFRKVGIPPIIALLSLASIIIVVLIKVILDKYYFL 63
Qy 62 CGQPLHFI PRKQLCDGELDCPLGEDEEHCVKSPFEGPAPAVRLSKDRSTLQVLD SATGNW 121
Db 64 CGQPLHFI PRKQLCDGELDCPLGEDEEHCVKSPFEGPAPAVRLSKDRSTLQVLD SATGNW 123
Qy 122 FSACPDNFTALAEATACROMGYSSKPTFRAVEIGPDQDLVDVEITENSQELRMNSSGPC 181
Db 124 FSACPDNFTALAEATACROMGYSSKPTFRAVEIGPDQDLVDVEITENSQELRMNSSGPC 183
Qy 182 LSGSLVSLHCLACGKSLKTPRVVGGEEASVDSWPQVSIQYDKQHVCGSILDPHWLTA 241
Db 184 LSGSLVSLHCLACGKSLKTPRVVGGEEASVDSWPQVSIQYDKQHVCGSILDPHWLTA 243
Qy 242 AHCPRKHTDVFNMKVRAGSKLGSFSPSLAVAKIIIEFNPMYPKNDIALMKLOPLTFS 301
Db 244 AHCPRKHTDVFNMKVRAGSKLGSFSPSLAVAKIIIEFNPMYPKNDIALMKLOPLTFS 303
Qy 302 GTVRPCLPFDEBLTPATPLMIIGWFTKQNGKMSDILLQASVQVIDSTRCNADDAVQ 361
Db 244 AHCPRKHTDVFNMKVRAGSKLGSFSPSLAVAKIIIEFNPMYPKNDIALMKLOPLTFS 303
Qy 302 GTVRPCLPFDEBLTPATPLMIIGWFTKQNGKMSDILLQASVQVIDSTRCNADDAVQ 361
Db 304 GTVRPCLPFDEBLTPATPLMIIGWFTKQNGKMSDILLQASVQVIDSTRCNADDAVQ 363
Qy 362 GEVTEKMWKAGIPGGVDTCOGDSGGPLMYQSDQHWVGVIVSWGCGGSPSTPGVYTKVS 421
Db 364 GEVTEKMWKAGIPGGVDTCOGDSGGPLMYQSDQHWVGVIVSWGCGGSPSTPGVYTKVS 423
Qy 422 AYLNWIYNVWKAEL 435
Db 424 AYLNWIYNVWKAEL 437

RESULT 15

US-10-473-127-1586
; Sequence 1586, Application US/10473127
; Publication No. US20040236091A1
; GENERAL INFORMATION:
; APPLICANT: Zycos Inc.

RESULT 14
US-10-473-127-1581
; Sequence 1581, Application US/10473127
; Publication No. US20040236091A1
; GENERAL INFORMATION:
; APPLICANT: Zycos Inc.
; TITLE OF INVENTION: TRANSLATIONAL PROFILING
; FILE REFERENCE: 08191-026W01
; CURRENT APPLICATION NUMBER: US/10/473,127
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 60/279,495
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: 60/292,544
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 60/310,801
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/326,370
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: 60/336,780
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/358,985
; PRIOR FILING DATE: 2002-02-20
; NUMBER OF SEQ ID NOS: 2041
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 1581
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-473-127-1581

Query Match 99.8%; Score 2337; DB 5; Length 437;
Best Local Similarity 100.0%; Pred. No. 5.9e-219;
Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 DPDSQPLNSLDVKLPKPRIPMETFRKVGIPPIIALLSLASIIIVVLIKVILDKYYFL 61
Db 4 DPDSQPLNSLDVKLPKPRIPMETFRKVGIPPIIALLSLASIIIVVLIKVILDKYYFL 63
Qy 62 CGQPLHFI PRKQLCDGELDCPLGEDEEHCVKSPFEGPAPAVRLSKDRSTLQVLD SATGNW 121
Db 64 CGQPLHFI PRKQLCDGELDCPLGEDEEHCVKSPFEGPAPAVRLSKDRSTLQVLD SATGNW 123
Qy 122 FSACPDNFTALAEATACROMGYSSKPTFRAVEIGPDQDLVDVEITENSQELRMNSSGPC 181
Db 124 FSACPDNFTALAEATACROMGYSSKPTFRAVEIGPDQDLVDVEITENSQELRMNSSGPC 183
Qy 182 LSGSLVSLHCLACGKSLKTPRVVGGEEASVDSWPQVSIQYDKQHVCGSILDPHWLTA 241
Db 184 LSGSLVSLHCLACGKSLKTPRVVGGEEASVDSWPQVSIQYDKQHVCGSILDPHWLTA 243
Qy 242 AHCPRKHTDVFNMKVRAGSKLGSFSPSLAVAKIIIEFNPMYPKNDIALMKLOPLTFS 301
Db 244 AHCPRKHTDVFNMKVRAGSKLGSFSPSLAVAKIIIEFNPMYPKNDIALMKLOPLTFS 303
Qy 302 GTVRPCLPFDEBLTPATPLMIIGWFTKQNGKMSDILLQASVQVIDSTRCNADDAVQ 361
Db 304 GTVRPCLPFDEBLTPATPLMIIGWFTKQNGKMSDILLQASVQVIDSTRCNADDAVQ 363
Qy 362 GEVTEKMWKAGIPGGVDTCOGDSGGPLMYQSDQHWVGVIVSWGCGGSPSTPGVYTKVS 421
Db 364 GEVTEKMWKAGIPGGVDTCOGDSGGPLMYQSDQHWVGVIVSWGCGGSPSTPGVYTKVS 423
Qy 422 AYLNWIYNVWKAEL 435
Db 424 AYLNWIYNVWKAEL 437

; TITLE OF INVENTION: TRANSLATIONAL PROFILING
; FILE REFERENCE: 08191-026W01
; CURRENT APPLICATION NUMBER: US/10/473,127
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 60/279,495
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: 60/292,544
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 60/310,801
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/326,370
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: 60/336,780
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/358,985
; PRIOR FILING DATE: 2002-02-20
; NUMBER OF SEQ ID NOS: 2041
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1586
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-473-127-1586

Query Match	99.8%	Score 2337;	DB 5;	Length 437;
Best Local Similarity	100.0%;	Pred. No. 5.9e-219;		
Matches 434;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	2	DPDSQPLNSLDVKPLRPKIPMETFRKVGPIIIALLSLASIIIVVLIKVILDKYYFL	61
Db	4	DPDSQPLNSLDVKPLRPKIPMETFRKVGPIIIALLSLASIIIVVLIKVILDKYYFL	63
Qy	62	CGQPLHFIIPKQLCDGELDCPLGEDEHEHCVKSPFEGPAVAVRLSKDRSTLQVLDSATGNW	121
Db	64	CGQPLHFIIPKQLCDGELDCPLGEDEHEHCVKSPFEGPAVAVRLSKDRSTLQVLDSATGNW	123
Qy	122	FSACFDNTEALAEATACRQMGYSKPTFRAVEIGDPDQDLDVVEITENSQELMRNMSGPC	181
Db	124	FSACFDNTEALAEATACRQMGYSKPTFRAVEIGDPDQDLDVVEITENSQELMRNMSGPC	183
Qy	182	LSGSLVSLHCLACGSKLTTPRVVGEEASVDSWPQVSIQYDKQHVCGGSIIDPHWVLT	241
Db	184	LSGSLVSLHCLACGSKLTTPRVVGEEASVDSWPQVSIQYDKQHVCGGSIIDPHWVLT	243
Qy	242	AHCFRKHTDVFNWKVRAGSDKLGSPSLAVAKIIIEFNPMYPKDNDIALMKLPPLTFS	301
Db	244	AHCFRKHTDVFNWKVRAGSDKLGSPSLAVAKIIIEFNPMYPKDNDIALMKLPPLTFS	303
Qy	302	GTVRPICLPFDEELTPATPLWIIWGFTKQNGKMSDILLOASVQVIDSTRCNADDAYQ	361
Db	304	GTVRPICLPFDEELTPATPLWIIWGFTKQNGKMSDILLOASVQVIDSTRCNADDAYQ	363
Qy	362	GEVTEKMMKAGIPGEGVDTCQDGGPLMYQSDQHHVVGIVSWGYCGGSPSTPGVYTKVS	421
Db	364	GEVTEKMMKAGIPGEGVDTCQDGGPLMYQSDQHHVVGIVSWGYCGGSPSTPGVYTKVS	423
Qy	422	AYLNIYINWKAEL 435	
Db	424	AYLNIYINWKAEL 437	

Search completed: December 5, 2005, 14:07:34
Job time : 168 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 6, 2005, 20:06:35 : Search time 236 Seconds
(without alignments)
3276.438 Million cell updates/sec

Title: US-10-803-530-2

Perfect score: 2342

Sequence: 1 MDPDSQDPLNSLDVKPLRKP.....VYTKVSALNWIYNWKAEK 435

Scoring table:

BLOSUM62	Xgapop 10.0	Xgapext 0.5
-G	Xgapop 10.0	Xgapext 0.5
Ygapop 6.0	Ygapext 7.0	
Delop 6.0	Delext 7.0	

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool.p/US10803530/runat_05122005_083216_2147/app.query.fasta_1.583
-DB=Issued Patents NA -OPMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -LOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10803530@cgn_1.1.290 @runat_05122005_083216_2147 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=10 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:**

1:	/cgn2_6/ptodata/1/ina/1 COMB.seq.*
2:	/cgn2_6/ptodata/1/ina/5 COMB.seq.*
3:	/cgn2_6/ptodata/1/ina/6A COMB.seq.*
4:	/cgn2_6/ptodata/1/ina/6B COMB.seq.*
5:	/cgn2_6/ptodata/1/ina/H COMB.seq.*
6:	/cgn2_6/ptodata/1/ina/PTUS COMB.seq.*
7:	/cgn2_6/ptodata/1/ina/pp COMB.seq.*
8:	/cgn2_6/ptodata/1/ina/RE COMB.seq.*
9:	/cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2338	99.8	2038	3	US-09-008-271A-18
2	2338	99.8	2038	3	US-09-968-415-18
3	2337	99.8	2081	3	US-09-851-588-7
4	2335	99.7	2121	3	US-09-607-745-1
5	2324	99.2	2079	3	US-09-656-002-1
6	2324	99.2	2079	3	US-09-851-588-5
7	2297.5	98.1	2063	3	US-10-012-231A-274
8	2297.5	98.1	2063	3	US-10-015-389A-274
9	2297.5	98.1	2063	3	US-10-006-768A-274

10	2297.5	98.1	2063	3	US-10-015-671A-274
11	2297.5	98.1	2063	3	US-10-015-333A-274
12	2297.5	98.1	2063	3	US-10-011-833A-274
13	2297.5	98.1	2063	3	US-10-006-041A-274
14	2297.5	98.1	2063	3	US-10-012-064A-274
15	1259	53.8	1189	3	US-09-607-745-8
16	698.5	29.8	1434	3	US-10-177-661-1
17	697.5	29.8	1748	3	US-09-879-792-11
18	683	29.2	1341	3	US-10-177-661-3
19	677.5	28.9	2479	3	US-09-342-749-29
20	677.5	28.9	2479	3	US-09-691-840-29
21	677.5	28.9	2479	3	US-09-685-166A-894
22	677.5	28.9	2479	3	US-09-679-426-894
23	677.5	28.9	2479	3	US-09-759-143-894
24	677.5	28.9	2479	3	US-10-012-896-894
25	676.5	28.9	1479	3	US-09-342-749-1
26	676.5	28.9	1479	3	US-09-691-840-1
27	676.5	28.9	2448	3	US-09-949-016-5203
28	675.5	28.8	1476	3	US-09-759-143-931
29	675.5	28.8	1476	3	US-10-012-896-931
30	675.5	28.8	1479	3	US-09-759-143-930
31	675.5	28.8	1479	3	US-10-012-896-930
32	675.5	28.8	3245	3	US-09-759-143-929
33	675.5	28.8	3245	3	US-10-012-896-929
34	663.5	28.3	3443	3	US-09-807-201-8
35	660.5	28.2	2440	3	US-09-949-016-5210
36	660.5	28.2	2440	3	US-09-949-016-5211
37	660.5	28.2	2440	3	US-09-949-016-5212
38	658	28.1	2412	3	US-09-999-833A-68
39	658	28.1	2412	3	US-10-020-445A-68
40	650.5	27.8	2413	3	US-09-518-046-1
41	650.5	27.8	2413	3	US-09-650-371-1
42	628.5	26.8	2070	3	US-09-856-371A-7
43	627	26.8	2544	3	US-09-518-046-3
44	627	26.8	2544	3	US-09-650-371-3
45	622.5	26.6	1230	3	US-09-879-792-35

SEQUENCE 274, Appl
SEQUENCE 274, Appl
SEQUENCE 274, Appl
SEQUENCE 274, Appl
SEQUENCE 274, Appl
SEQUENCE 274, Appl
SEQUENCE 1, Appl
SEQUENCE 1, Appl
SEQUENCE 11, Appl
SEQUENCE 3, Appl
SEQUENCE 29, Appl
SEQUENCE 29, Appl
SEQUENCE 894, Appl
SEQUENCE 894, Appl
SEQUENCE 894, Appl
SEQUENCE 894, Appl
SEQUENCE 1, Appl
SEQUENCE 1, Appl
SEQUENCE 5203, Ap
SEQUENCE 931, Appl
SEQUENCE 931, Appl
SEQUENCE 930, Appl
SEQUENCE 929, Appl
SEQUENCE 929, Appl
SEQUENCE 8, Appl
SEQUENCE 5210, Ap
SEQUENCE 5212, Ap
SEQUENCE 68, Appl
SEQUENCE 68, Appl
SEQUENCE 1, Appl
SEQUENCE 1, Appl
SEQUENCE 7, Appl
SEQUENCE 3, Appl
SEQUENCE 35, Appl

US-09-008-271A-18
; Sequence 18, Application US/09008271A
; Patent No. 6203979
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; Hillman, Jennifer L.
; Yue, Henry
; Guegler, Karl J.
; Corley, Neil C.
; Tang, Tom Y.
; Shah, Purvi
; TITLE OF INVENTION: HUMAN PROTEASE MOLECULES
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION NUMBER: US/09/008,271A
; FILING DATE: 16-Jan-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:

```

; NAME: Mohan-Peterson, Sheela
; REGISTRATION NUMBER: 41,201
; REFERENCE/DOCKET NUMBER: PP-0458 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2038 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: COLNOT13
; CLONE: 1337018
; SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-09-008-271A-18

Alignment Scores:
Pred. No.:      8,646-237      Length:      2038
Score:          2338.00      Matches:    434
Percent Similarity: 100.00%      Conservative: 1
Best Local Similarity: 99.77%      Mismatches: 0
Query Match:      99.83%      Indels:    0
Dbs:              3          Gaps:      0

US-10-803-530-2 (1-435) x US-09-008-271A-18 (1-2038)

QY      1 MetAspProAspSerAspGluProLeuAsnSerLeuAspValLysProLeuArgLysPro 20
Db      200 ATGGATCTGACAGTGAATCAACCTCTGAACAGCTCGATGTCAAACCCCTCGCGAAACCC 259

QY      21 ArgIleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSer 40
Db      260 CGTATCCCATGGAGACCTTCAGAAAGGTGGGATCCCATCATCATAGCACTACTGAGC 319

QY      41 LeuAlaSerIleIleIleValValValLysValIleLeuAspLysTyrTyrPhe 60
Db      320 CTGGCGAGTATCATATTGTGGTTGTCTCATCAAGGTGATTCTGGATAAATACTACTTC 379

QY      61 LeuCysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAsp 80
Db      380 CTCTGGGGGAGCCCTCTCCATTCATCCGAGGAGCAGCTGTGTGACGGAGCTGGAC 439

QY      81 CysProLeuGlyGluAspGluGluHisCysValLysSerPheProGluGlyProAlaVal 100
Db      440 TGTCCCTTGGGGGAGGACGAGGACACTGTGTCAAGAGCTTCCCGAAGGCGCTGCAAGTG 499

QY      101 AlaValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsn 120
Db      500 GCAGTCCGCTCTCCAAAGGACCGATCCACACTGCAGGTGTGGACTCGGCCACAGGGGAC 559

QY      121 TrpPheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysAtgGln 140
Db      560 TGGTTCTCTGCTCTTTTCGCAACTTCACAGAGCTCTCGCTGAGACAGCTGTAGGCAG 619

QY      141 MetGlyTyrSerSerLysProThrPheArgAlaValGluLeuGlyProAspGlnAspLeu 160
Db      620 ATGGGCTACAGCAGCAAAACCCACTTTCAGAGCTGTGAGATTGGCCACAGCAGGATCTG 679

QY      161 AspValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyPro 180
Db      680 GATGTTGTTGNAATCAGAAAACAGCCAGAGCTTCGCATGCGGAATCAAGTGGGCCC 739

QY      181 CysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThr 200
Db      740 TGTCTCTCAGGCTCCCTGGTCTCCCTGCACTGTCTTCCCTGTGGGAGACCTGAAGACC 799

QY      201 ProArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIle 220
Db      800 CCCCCTGTGGTGGGAGGAGGCGCTCTGTGATTCTTTGGCTTTGGCAGGTGAGCATC 859

QY      221 GlnTyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThr 240

```

```

Db      860 CAGTACGACAAACAGCACGCTCTGTGGAGGAGCATCTCGACCCCACTGGTCTCTCACG 919
QY      241 AlaAlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySer 260
Db      920 GCAGCCCACTGCTTCAGGAACATACCGATGTGTCACTGGAGGTGCGGAGGGCTCA 979
QY      261 AspLysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsn 280
Db      980 GACAAACTGGGCAGCTTCCCATCCCTGGCTGTGGCAAGATCATCATTTGAATTCAAC 1039
QY      281 ProMetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPhe 300
Db      1040 CCCATGTACCCCAAGACATGATGCGCCCTCATGAAGCTGCAGTTCACCACTCATTTC 1099
QY      301 SerGlyThrValArgProIleCysLeuProPheAspGluGluLeuThrProAlaThr 320
Db      1100 TCAGGCACAGTCAGGCCCATCTCTCTGCTCTTTTGTATGAGGAGCTCACTCCAGCCACC 1159
QY      321 ProLeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIle 340
Db      1160 CCACCTCTGGATCATTCGATGGGCTTTACGAAGCAGATGGAGGAAGATGTCTGACATA 1219
QY      341 LeuLeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaTyr 360
Db      1220 CTGCTGCAGGCGTCAGTCCAGGTCAATTGACGACACAGGTGCAATGCAGACGATGCGTAC 1279
QY      361 GlnGlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThr 380
Db      1280 CAGGGGGAAGTCAACCAGAAAGATGATGTGTGAGGATCCCGGAAGGGGTGTGGACACC 1339
QY      381 CysGlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGly 400
Db      1340 TGCAGGGTGACAGTGGTGGGCCCCGTGATGTACCAATCTGCACAGTGGCATGTGGTGGC 1399
QY      401 IleValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysVal 420
Db      1400 ATCGTTAGCTGGGCTATGCTGCGGGGCGCCGAGCACCCAGGAGTATACACCAAGGTC 1459
QY      421 SerAlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db      1460 TCAGCCATATCTCAACTGGATCTACAATGTCTGGAAGGCTGAGCTG 1504

RESULT 2
US-09-968-415-18
; Sequence 18, Application US/09968415
; Patent No. 6855811
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; Hillman, Jennifer L.
; Yue, Henry
; Guegler, Karl J.
; Corley, Neil C.
; Tang, Tom Y.
; Shah, Purvi
; TITLE OF INVENTION: HUMAN PROTEASE MOLECULES
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESS: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/968,415
; FILING DATE: 26-Sep-2001
; PRIOR APPLICATION DATA:

```

```
; APPLICATION NUMBER: 09/659,151
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Mohan-Peterson, Sheela
; REGISTRATION NUMBER: 41,201
; REFERENCE/DOCKET NUMBER: PF-0458 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2038 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: COLNOT13
; CLONE: 1337018
; SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-09-968-415-18

Alignment Scores:
Pred. No.:      8,648-237      Length:      2038
Score:          2338.00      Matches:      434
Percent Similarity: 100.00%      Conservative: 1
Best Local Similarity: 99.77%      Mismatches: 0
Query Match:      99.83%      Indels:      0
DB:              3          Gaps:      0

US-10-803-530-2 (1-435) x US-09-968-415-18 (1-2038)

QY 1 MetAspProAspSerAspGlnProLeuAenSerLeuAenValLysProLeuArgLysPro 20
DB 200 ATGGATCTCTGACAGTATCAACCTCTGAACAGCCTCGATGTCAAAACCCCTGCGCAACCC 259
QY 21 ArgIleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuSer 40
DB 260 CGTATCCCATGGAGACCTTCAGAAAGGTGGGGATCCCCCATCATATGACACTACTGAGC 319
QY 41 LeuAlaSerIleIleIleValValValLeuIleLysValIleLeuAspLysTyrTyrPhe 60
DB 320 CTGGCGAGTATCATATTGTGGTGTCTCATCAAGGTGATTCTGGATAAATACTACTTC 379
QY 61 LeuCysGlyGlnProLeuHisPheLeuProArgLysGlnLeuLysAspGlyGluLeuAsp 80
DB 380 CTCCTGGGCGAGCCTCTCCACTTCATCCCGAGAGCAGCTGTGTGACGAGAGCTGGAC 439
QY 81 CysProLeuGlyGluAspGluHisCysValLysSerPheProGluGlyProAlaVal 100
DB 440 TGTCCCTTGGGGGAGGACGAGAGCAGTGTGTCAAGAGCTTCCCGAAGGGCTCGAGTG 499
QY 101 AlaValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAen 120
DB 500 GCGATCCCGCTCTCAAGAGCCAGATCCACACTGCAAGTGTCTGCTCGGCCCAAGGGAAC 559
QY 121 TrpPheSerAlaCysPheAspAenPheThrGluAlaLeuAlaGluThrAlaCysArgGln 140
DB 560 TGGTCTCTGCCTGTTTCGACAACTTCACAGAACTCTCGCTGAGACACCTGTAGGCAG 619
QY 141 MetGlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeu 160
DB 620 ATGGGTACAGCAGCAAAACCCACTTTTCAGAGCTGTGGAGATTGGCCCGACAGCAGGATCTG 679
QY 161 AspValValGluIleThrGluAenSerGlnGluLeuArgMetArgAenSerSerGlyPro 180
DB 680 GATGTGTGTTGAATAACAGAAACAGCCAGGAGCTTCGCATCGGAACCTCAAGTGGGCC 739
QY 181 CysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThr 200
DB 740 TGTCTCTAGGCTCCCTGCTCTCCCTGCACTGTCTTGCTGTGGGAGAGCCTGAAGACC 799
QY 201 ProArgValValGlyGlyGluGluAlaSerValAspSerTyrProTyrGlnValSerIle 220
```

```
DB 800 CCCCCTGTGGTGGTGGGAGAGGCGCTCTGTGGATTCTTGGCCCTTGGCAGGTGAGCATC 859
QY 221 GlnTyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTyrValLeuThr 240
DB 860 CAGTACGACAAACAGCAGCTGTGTGGAGGAGCATCTCTGGACCCCACTGGGTCTCAGC 919
QY 241 AlaAlaHisCysPheArgLysHisThrAspValPheAenTyrLysValArgAlaGlySer 260
DB 920 CGAGCCCACTGCTTCAGGAAACATACCGATGTTCAACTGGAGGTGCGGCGAGGCTCA 979
QY 261 AspLysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleIleGluPheAen 280
DB 980 GACAACTGGGCGAGCTTCCCATCTGCTGTGGCCAAAGATCATCATTAATTAAC 1039
QY 281 ProMetTyrProLysAspAenAspIleAlaLeuMetLysLeuGlnPheProLeuThrPhe 300
DB 1040 CCCATGTACCCCAAGACAATGACATCGCCCTCATGAAGCTGCAGTCCCACTCACTTTC 1099
QY 301 SerGlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThr 320
DB 1100 TCAGGCACAGTCAGGCCCATCTGTCTGCCCTTCTTTGATGAGGAGCTCACTCCAGCCACC 1159
QY 321 ProLeuTyrIleIleGlyTyrGlyPheThrLysGlnAenGlyGlyLysMetSerAspIle 340
DB 1160 CCACCTGTGATCATTTGGATGGGCTTTACGAAGCAGATGGAGGGAAGATGTCTGACATA 1219
QY 341 LeuLeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaTyr 360
DB 1220 CTGCTGCAGGCGTCAGTCCAGTCAATTCACAGCACCGGTGCAATGCAGACGATCGTAC 1279
QY 361 GlnGlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThr 380
DB 1280 CAGGGGGAAGTCAACGAGAAGATGATGTGCAGGCAATCCCGAAGGGGTGTGGACACC 1339
QY 381 CysGlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTyrHisValValGly 400
DB 1340 TGCCAGGCTGACAGTGTGGGCCCCCTGATGATCCCAATCTGACCACTGGCATGTGTGGGC 1399
QY 401 IleValSerTyrGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysVal 420
DB 1400 ATCGTTAGCTGGGCTATGGCTGCGGGGCGCCGAGCACCCAGGAGTATACACCAAGGTC 1459
QY 421 SerAlaTyrLeuAenTyrIleTyrAsnValTyrLysAlaGluLeu 435
DB 1460 TCAGCTATCTCACTGGATCTACAATGTCTGGAAGGCTGAGCTG 1504

RESULT 3
; Sequence 7, Application US/09851588
; Patent No. 6682890
; GENERAL INFORMATION:
; APPLICANT: Mack, David
; APPLICANT: Gish, Kurt C.
; APPLICANT: Wilson, Keith E.
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSING COLORECTAL CANCER, COMPOSITIONS, AND
; TITLE OF INVENTION: OF SCREENING FOR COLORECTAL CANCER MODULATORS
; FILE REFERENCE: A-68829-1/DJB/JJD/AMS
; CURRENT APPLICATION NUMBER: US/09/851,588
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 09/642,252
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: US 09/656,002
; PRIOR FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 7
; LENGTH: 2081
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (215)..(1528)
; OTHER INFORMATION:
```

US-09-851-588-7

Alignment Scores:

Pred. No.: 1,14e-236 Length: 2081
Score: 2337.00 Matches: 434
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.79% Indels: 0
DB: 3 Gaps: 0

US-10-803-530-2 (1-435) x US-09-851-588-7 (1-2081)

QY 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysProArg 21
DB 224 GATCCTGACAGTGATCAACCTCTGAACAGCCTCGATGTCAAAACCCCTCGCAAAACCCGT 283
QY 22 IleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSerLeu 41
DB 284 ATCCCATGGAGACCTTCAGAAAGGTGGGATCCCATCATCATAGCACTACTGAGCCTG 343
QY 42 AlaSerIleIleIleValValLeuIleLysValIleLeuAspLysTyrTyrPheLeu 61
DB 344 GCGAGTATCATATTGTGTTCTCATCAAGGTGATTCTGGATAATACTACTTCTC 403
QY 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys 81
DB 404 TCGGGGAGCCCTCTCCACTTCATCCGAGGAGCAGCTGTGTGACGAGAGCTGGACTGT 463
QY 82 ProLeuGlyGluAspGluGluHisCysValLysSerPheProGluGlyProAlaValAla 101
DB 464 CCTCTGGGGAGGACGAGAGCACTGTGTCAAGAGCTTCCCGAAGGCCCTGCGAGTGGA 523
QY 102 ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp 121
DB 524 GTCCGCCCTCTCCAGGACCGATCCACACTGCAGGTGCTGGACTCGGCCACAGGGAAGTGG 583
QY 122 PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
DB 584 TTCTCTGCTGTTTCGACACTTCACAGAACTCTCGCTGAGACAGCTGTGAGCAGATG 643
QY 142 GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
DB 644 GGCTACAGCAGCAAAACCCACTTTCAGAGCTGTGAGATTCGCCAGACCAAGATCTGGAT 703
QY 162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
DB 704 GTGTGTGAATTCACAGAAACAGCAGGAGCTTCGCATGCGAACTCAAGTGGGCCCTGT 763
QY 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro 201
DB 764 CTCTCAGGCTCCCTGGTCTCCCTGCATGCTTGCTTGTTGGTGGAGAGCCTGAAGACCCC 823
QY 202 ArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
DB 824 CGTGTGTGGTGGGAGGAGCCCTCTGTGATTCTTGGCTTGGCAGGTGAGCATCCAG 883
QY 222 TyrAspLysGlnHisValCysGlySerIleLeuAspProHisTrpValLeuThrAla 241
DB 884 TACGACAAACAGCAGCTGTGGAGGAGGATCTCGAACCCCACTGGGTCTCTCAGGCA 943
QY 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
DB 944 GCCCACTGCTTCAGAAACATACCGATGCTTCACTGGAGGTGGGGAGGCTCAGAC 1003
QY 262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsnPro 281
DB 1004 AAATGGGAGCTTCCATCCCTGGCTGTGGCCAAAGATCATCATTTGAATTCACACCC 1063
QY 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
DB 1064 ATGTACCCCAAGACAATGATCGCCCTCATGAAGCTGAGTTCACACTCACTTTCTCA 1123
QY 302 GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro 321

DB 1124 GGCACAGTCAGGCCCATCTGTCTGCCCTTTTTCATGAGGAGCTCACTCCAGCCACCCCA 1183
QY 322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyLysMetSerAspIleLeu 341
DB 1184 CTCTGGATCATTTGGATGGGCTTTTACGAAGCAGAAATGGAGGAAGATGTCTGACATACTG 1243
QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGln 361
DB 1244 CTGAGGGCTCAGTCCAGGTCAATTGACAGCACACGGTGCATGACAGACGATGCGTACCAG 1303
QY 362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys 381
DB 1304 GGGGAAGTCACCCGAGAAGATGATGTGTGAGGACATCCCGAAGGGGTGTGGACACTGC 1363
QY 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle 401
DB 1364 CAGGGTGACAGTGTGGGCCCTGTATGTACCAATCTGACCAGTGGCATGTGGTGGGCATC 1423
QY 402 ValSerTrpGlyTyrGlyCysGlyProSerThrProGlyValTyrThrLysValSer 421
DB 1424 GTTAGCTGGGCTATGCTCGGGGGCCGAGCACCCAGAGTATACACCAAGGTCTCA 1483
QY 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
DB 1484 GCCTATCTCAACTGGATCTACAATGTCTGGAAGGCTGAGCTG 1525

RESULT 4

US-09-607-745-1
; Sequence 1, Application US/09607745
; Patent No. 6750034
; GENERAL INFORMATION:
; APPLICANT: Darrow, Andrew L
; APPLICANT: Qi, Jain-shen
; APPLICANT: Andrade-Gordon, Patricia
; TITLE OF INVENTION: DNA encoding human serine protease D-G
; FILE REFERENCE: ORT-1273
; CURRENT APPLICATION NUMBER: US/09/607,745
; CURRENT FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2121
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-607-745-1

Alignment Scores:

Pred. No.: 1,91e-236 Length: 2121
Score: 2335.00 Matches: 434
Percent Similarity: 99.77% Conservative: 0
Best Local Similarity: 99.77% Mismatches: 1
Query Match: 99.70% Indels: 0
DB: 3 Gaps: 0

US-10-803-530-2 (1-435) x US-09-607-745-1 (1-2121)

QY 1 MetAspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysPro 20
DB 277 ATGGATCTCTGACAGTGATCAACCTCTGAACAGCCTCGATGTCAAAACCCCTGCGCAACCC 336
QY 21 ArgIleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSer 40
DB 337 CGTATCCCATGGAGACCTTCAGAAAGGTGGGATCCCATCATCATGACATCATGAGC 396
QY 41 LeuAlaSerIleIleIleValValLeuIleLysValIleLeuAspLysTyrTrpPhe 60
DB 397 CTGGGAGTATCATATTGTGTTGTCTCATCAAGGTGATTCTGGATAAATACTACTTC 456
QY 61 LeuCysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAsp 80
DB 457 CTCTCGGGGAGCCTCTCCACTTTCATCCCGAAGAGCAGCTGTGTGACGGAGCTGGAC 516

```
QY 81 CysProLeuGlyGluAspGluHisCysValLysSerPheProGluGlyProAlaVal 100
Db 517 TGTCCCTTGGGAGAGCAGGAGACCTGTGTCAAGAGCTTCCCGAAGGGCTGCAGTG 576
QY 101 AlaValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsn 120
Db 577 GCAGTCCGCTCTCCAAAGACCGATCCACACTGCAGTGTCTGGACTCGGCCACAGGGAAC 636
QY 121 TrpPheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGln 140
Db 637 TGGTTCCTCTGCTGTGTTCACAACTTTCACAGAAGCTCTCGCTGAGACAGCCTGTAGGCAG 696
QY 141 MetGlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeu 160
Db 697 ATGGGCTACAGCAGCAAAACCCACTTTCAGAGCTGTGGAGATTGGCCACAGCAGGATCTG 756
QY 161 AspValValGluIleThrGluAsnSerGlnLeuArgMetArgAsnSerSerGlyPro 180
Db 757 GATGTTGTTGAAATCACAGAAACAGCCAGGAGCTTCGCATCGGGAACCTCAAGTGGGCC 816
QY 181 CysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThr 200
Db 817 TGTCTCTCAGGCTCCCTGCTCCCTGCACCTGTCTTGCCTGTGGGAAGAGCCTGAAGACC 876
QY 201 ProArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIle 220
Db 877 CCCCCTGTGTGGTGGGAGAGGCTCTGTGGATCTTGGCCTTGGCAGGTTCAGCATC 936
QY 221 GlnTyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThr 240
Db 937 CAGTACGACAAACAGCACCTCTGTGGAGGAGCATCTCTGACCCCTCAGTGGTCTCAGC 996
QY 241 AlaAlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySer 260
Db 997 GCAGCCCACTGCTTCAGGAACATACCGATGTGTTCAACTGGAAGTGGCGGCGAGCTCA 1056
QY 261 AspLysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsn 280
Db 1057 GACAACTGGGAGCTTCCATCTCCCTGGCTGTGGCCAAAGATCATCATCAATGAATCAAC 1116
QY 281 ProMetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPhe 300
Db 1117 CCATGTACCCCAAGACATGACATCGCCCTCATGAAGCTGCAGTTCACACTCATCTTC 1176
QY 301 SerGlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThr 320
Db 1177 TCAGGCACAGTCAGGCCCATCTGTCTGCCCTTCTTTGATGAGGAGCTCACTCCAGCCACC 1236
QY 321 ProLeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIle 340
Db 1237 CCACCTCTGGATCATTTGGATGGGGCTTTACGAAGCAGAAATGGAGGGAAGATGTCTGACA 1296
QY 341 LeuLeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyr 360
Db 1297 CTGCTGCAGCGTCACTCAGGTCAATGACGACACGGTGCATGCAATCAGACAGATGGGTAC 1356
QY 361 GlnGlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThr 380
Db 1357 CTGGGGGAAGTCACCCAGGAAGATGATGTGTGCAGGCATCCCGAAGGGGTGTGGACACC 1416
QY 381 CysGlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGly 400
Db 1417 TGCAGGGTGACAGTGTGGGCCCTTGATGTACCAATCTGACCAGTGGCATGTGTGGTGGGC 1476
QY 401 IleValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysVal 420
Db 1477 ATCGTTAGCTGGGGCTATGGCTCGGGGGGCCCGGAGCACCCTCCAGGGGTATACCAAGGTC 1536
QY 421 SerAlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db 1537 TCAGCCATCTCACTGGATCTACATGTCTGGAGGCTGAGCTG 1581
```

RESULT 5

```
US-09-656-002-1
; Sequence 1, Application US/09656002
; Patent No. 6455668
; GENERAL INFORMATION:
; APPLICANT: Mack, David
; APPLICANT: Gish, Kurt
; APPLICANT: Wilson, Keith
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSING COLORECTAL CANCER, COMPOSITIONS, AND
; FILE REFERENCE: A-69108/DJB/JJD/AMS
; CURRENT APPLICATION NUMBER: US/09/656,002
; CURRENT FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: US 09/525,993
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: US 09/493,444
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: PCT/US 00/07044
; PRIOR FILING DATE: 2000-03-15
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 2079
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-656-002-1
Alignment Scores:
Pred. No.: 2,68e-235 Length: 2079
Score: 2324.00 Matches: 434
Percent Similarity: 99.77% Conservative: 0
Best Local Similarity: 99.77% Mismatches: 0
Query Match: 99.23% Indels: 1
DB: 3 Gaps: 0
US-10-803-530-2 (1-435) x US-09-656-002-1 (1-2079)
```

```
QY 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysProArg 21
Db 217 GATCTCGACAGTGATCAACCTCTGAACAGCCTCGATGTCAAAACCCCTCGCAACCCCTG 276
QY 22 IleProMetGluThrPheArgLys-ValGlyIleProIleIleIleAlaLeuSerLe 41
Db 277 ATCCCCATGGAGACCTTCAGAAAGTGTGGGATGCCCATCATCATAGCACTACTGACCT 336
QY 41 uAlaSerIleIleIleValValValLeuIleLysValIleLeuAspLysTyrTyrPheLe 61
Db 337 GCGAGATATCATTTGTGTGTCTCATCAAGGTGATTCCTGGATAAATACTACTTCTCT 396
QY 61 uCysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCy 81
Db 397 CTGCGGGCAGCCTCTCCACTTCCCGAGGAAGCAGCTGTGTGACGGAGAGCTGGACTG 456
QY 81 sProLeuGlyGluAspGluGluHisCysValLysSerPheProGluGlyProAlaValAl 101
Db 457 TCCCTTGGGGAGGACGAGGACATGTGTCAAGAGCTTCCCGAAGGGCCTTGACGTGGC 516
QY 101 aValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTr 121
Db 517 AGTCCGCTCTCCAAAGACCGATCCACACTGCAGGTGTGTGGACTCGGCCACAGGGAAC 576
QY 121 pPheSerAlaCysPheAspAsnPheThrGluAlaIleAlaGluThrAlaCysArgGlnMe 141
Db 577 GTTCTCTGCTCTTTTCGACAACTTCACAGAAGCTCTCGCTGAGACAGCCTGTAGGCAGAT 636
QY 141 tGlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAs 161
Db 637 GGGCTACAGCAACAAACCCACTTTTCAGAGCTGTGGAGATTGGCCACAGCAGATCTGGA 696
QY 161 pValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCy 181
Db 697 TGTGTTGAAATCACAGAAACAGCCAGGAGCTTCGATCGCGAAGCTCAAGTGGGCCCTG 756
QY 181 sLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPr 201
```

Db 757 TCTCTCAGGCTCCCTGGTCTCCCTGCACCTGTCTTGGTCTGGGAGAGCCTGAAGACCCC 816
QY 201 oArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleG1 221
Db 817 CCGTGTGGTGGTGGGAGAGAGCCCTCTGTGGATTCTTGGCCCTGGCAGTCAAGTCCA 876
QY 221 nTyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAl 241
Db 877 GTACGACAAACAGCAGCTGTGTGGAGGAGCATCTTGGACCCCACTGGGTCTCACGGC 936
QY 241 aAlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAs 261
Db 937 AGCCCACTCTTTCAGGAAACATACCATGTTTCAACTGGAAGGTGGCGCAGGCTCAGA 996
QY 261 pLysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleIleGluPheAsnPr 281
Db 997 CAAACTGGGAGCTTCCCACTCCCTGGCTGTGGCCAGATCATCATGTGAATTCACCC 1056
QY 281 oMetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSe 301
Db 1057 CATGTACCCCAAGACAATGACATCGCCCTCATGAAGCTGCAGTTCCCACTCACTTCTC 1116
QY 301 rGlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPr 321
Db 1117 AGGCACAGTCAGGCCCATCTGTCTGCCCTTCTTGTATGAGGAGCTCACTCCAGCCACCCC 1176
QY 321 oLeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLe 341
Db 1177 ACTCTGGATCAITGGATGGGCTTTACGAAGCAGAAATGGAGGAAGATGTCGTGACATACT 1236
QY 341 uLeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrgl 361
Db 1237 GCTGCAGGCGTCAGTCCAGGTCAATGACACACACGTCGAATGCAGACGATGGTACCA 1296
QY 361 nGlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCy 381
Db 1297 GGGGGAAGTCCAGAGAGATGATGTGCAGGATCCCGAAGGGGTGTGGACACCTG 1356
QY 381 sGlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyI1 401
Db 1357 CCAGGTCACAGTGTGTGGCCCTGATGATACCAATCTGACCAGTGGCATGTGTGGCAT 1416
QY 401 eValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValThrLysValSe 421
Db 1417 CTTAGCTGGGCTATGGCTCGGGGGCCGAGCACCCAGAGTATACACCAAGGTCTC 1476
QY 421 rAlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db 1477 AGCCTATCTCACTGGATCTACATGTCGTGAAGGCTGAGCTG 1519

RESULT 6

US-09-851-588-5
; Sequence 5, Application US/09851588
; Patent No. 6682890
; GENERAL INFORMATION:
; APPLICANT: Mack, David
; APPLICANT: Gish, Kurt C.
; APPLICANT: Wilson, Keith E.
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSING COLORECTAL CANCER, COMPOSITIONS, AND
; FILE OF INVENTION: OF SCREENING FOR COLORECTAL CANCER MODULATORS
; FILE REFERENCE: A-68829-1/DJB/JJD/AMS
; CURRENT APPLICATION NUMBER: US/09/851,588
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 09/642,252
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: US 09/656,002
; PRIOR FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 2079
; TYPE: DNA

; ORGANISM: Homo sapiens
US-09-851-588-5

Alignment Scores:
Pred. No.: 2,686-235 Length: 2079
Score: 2324.00 Matches: 434
Percent Similarity: 99.77% Conservative: 0
Best Local Similarity: 99.77% Mismatches: 0
Query Match: 99.23% Indels: 1
DB: 3 Gaps: 0

US-10-803-530-2 (1-435) x US-09-851-588-5 (1-2079)

QY 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysProArg 21
Db 217 GATCCTGACAGTGTATCAACCTCTGAACAGCCTCGATGTCAACCCCTGCGCAACCCCGT 276
QY 22 IleProMetGluThrPheArgLys-ValGlyIleProIleIleIleAlaLeuSerLe 41
Db 277 ATCCCATGGAGACCTTCAGAAAGTGTGGGATCCCCATCATATAGCCTACTGAGCCT 336
QY 41 uAlaSerIleIleIleValValValLeuIleLysValIleLeuAspLysTyrTyrPheLe 61
Db 337 GCGCAGTATCATATTGTGTCTCTCATCAAGGTGATTCTCGATAAATACTACTTCT 396
QY 61 uCysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCy 81
Db 397 CTGGGGAGCCCTCTCCACTTCATCCCGAGGAAGCAGCTGTGTGACGGAGAGCTGGACTG 456
QY 81 sProLeuGlyGluAspGluGluHisCysValLysSerPheProGluGlyProAlaValAl 101
Db 457 TCCCTTGGGGAGGAGCAGGAGCCTGTGTCAAGAGCTTCCCGAAGGGCTCGAGTGGC 516
QY 101 aValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTr 121
Db 517 AGTCCGCCTCTCCAAAGGACCGATCCACACTGTGACAGGTGTGGACTCGGCCACAGGAACTG 576
QY 121 pPheSerAlaCysPheAspAsnPherThrGluAlaLeuAlaGluThrAlaCysArgGlnMe 141
Db 577 GTTCTCTGGCTGTTCGACAACTTCACAGAACTCTCGCTGAGACAGCTGTAGGCAGAT 636
QY 141 tGlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAs 161
Db 637 GGGCTACAGCAGCAAAACCCACTTTCAGAGCTGTGGAGATTGGCCACAGACAGGATCTGA 696
QY 161 pValValGluIleThrGluAsnSerGlnLeuArgMetArgAsnSerSerGlyProCy 181
Db 697 TGTGTGTGAATACACAGAAACAGCCAGGAGCTTCGCATGCGGAACTCAAGTGGGCCCTG 756
QY 181 sLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPr 201
Db 757 TCTCTCAGGCTCCCTGGTCTCCCTGCACCTGTTCCTGTGGAAAGAGCTTGAAGACCCC 816
QY 201 oArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleG1 221
Db 817 CCGTGTGGTGGTGGGAGGAGCCCTCTGTGGATTCTTGGCCCTGGCAGTCAAGTCCA 876
QY 221 nTyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAl 241
Db 877 GTACGACAAACAGCAGCTGTGTGGAGGAGCATCTTGGACCCCACTGGGTCTCACGGC 936
QY 241 aAlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAs 261
Db 937 AGCCCACTGCTTCAGGAAACATACCATGTTTCAACTGGAAGGTGGCGCAGGCTCAGA 996
QY 261 pLysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleIleGluPheAsnPr 281
Db 997 CAAACTGGGAGCTTCCCACTCCCTGGCTGTGGCCAGATCATCATGTGAATTCACCC 1056
QY 281 oMetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSe 301
Db 1057 CATGTACCCCAAGACAATGACATCGCCCTCATGAGGCTGAGTTCCTCCACTCACTTCTC 1116

```
QY 301 rGlyThrValArgProIleCysLeuProPhePhePheGluGluLeuThrProAlaThrPr 321
Db 1117 AGGCACAGTCAGGCCCATCTGTCGCCCTCTTTGATGAGGAGCTCACTCCAGCCACCCC 1176
QY 321 oLeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLe 341
Db 1177 ACTCTGGATCATTTGGATGGGCTTTACGAAGCAGAAATGGAGGGAAGATGCTGTACATACT 1236
QY 341 uLeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaThrG1 361
Db 1237 GCTGAGGCGTCAGTCCAGGTCAATTGACAGCACACGGTCAATGACAGCGATGCGTACCA 1296
QY 361 nGlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCy 381
Db 1297 GGGGGAAGTCACCGAAGATGATGTGTCAGGCAATCCCGAAGGGGTGTGCACACCTG 1356
QY 381 sGlnGlyAspSerGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyI1 401
Db 1357 CCAGGCTGACAGTGTGGGCCCTCGATGTACCAATCTGACCAGTGGCATGTGTGGGCAT 1416
QY 401 eValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSe 421
Db 1417 COTTAGCTGGGCTATGGCTGGGGGGCCCGAGCACCCCGAGGATATACACCAAGGTCTC 1476
QY 421 rAlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db 1477 AGCCTATCTCAACTGATCTACAATGTCTGGAAGGCTGAGCTG 1519
```

RESULT 7

```
US-10-012-231A-274
; Sequence 274, Application US/10012231A
; Patent No. 6924355
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2830PIC23
; CURRENT APPLICATION NUMBER: US/10/012,231A
; PRIOR FILING DATE: 2002-06-10
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 274
; LENGTH: 2063
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-012-231A-274
```

```
Alignment Scores:
Pred. No.: 1,67e-232 Length: 2063
Score: 2297.50 Matches: 429
Percent Similarity: 98.85% Conservative: 0
Best Local Similarity: 98.85% Mismatches: 0
Query Match: 98.10% Indels: 5
DB: 3 Gaps: 1
```

US-10-803-530-2 (1-435) x US-10-012-231A-274 (1-2063)

```
QY 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysProArg 21
Db 219 GATCTTGACAGTGATCAACTCTGAACAGCCCTCGATGTCAAAACCCCTCGCGCAACCCCGT 278
```

```
QY 22 IleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSerLeu 41
Db 279 ATCCCCATGGAGACCTTCAGAAAGGTGGGGATCCCCATCATCATAGCACTACTAGACCTG 338
QY 42 AlaSerIleIleIleValValValLeuIleLysValIleLeuAspLysTyrTyrPheLeu 61
Db 339 GCGAGTATCATATTGTGTGTTCTCATCAAGGTGATTCTGGATAAATACTACTTCTCTC 398
QY 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys 81
Db 399 TCGGGCAGCCTCTCCACTTCATCCGAGGAAGCAGCTGTGTGACGAGAGCTGGAAGTGT 458
QY 82 ProLeuGlyGluAspGluGluHisCysValLysSerPheProGluGlyProAlaValAla 101
Db 459 CCCTTGGGGGAGGACGAGGACCTGTCTCAAGAGCTTCCCCGAAGGGCTCTCAGTGGCA 518
QY 102 ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp 121
Db 519 GTCCGCTCTCCAGGACCGATCCACACTGCAGGTGCTGGACTCGGCCACAGGAACCTGG 578
QY 122 PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
Db 579 TTCTCTGCTGTTTCGACAACTTCACGAAGCTCTCGCTGACAGACGCTGTAGGCAGATG 638
QY 142 GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
Db 639 GGCTACAGC-----AGAGCTGTGGAGATTGGCCCGACAGACGAGATCTGGAT 683
QY 162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
Db 684 GTTGTGTAATCAGCAAAAACAGCCAGGAGCTTCGCATGCGGAACCTCAAGTGGGCGCTGT 743
QY 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro 201
Db 744 CTCTCAGGCTCCCTGGTCTCCCTGCACTGTCTTGGCTGTGGAGAGAGCCCTGAAGACCCC 803
QY 202 ArgValValGlyGlyGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
Db 804 COTGTGGTGGGTGGGGAGGAGGCTCTGTGGATTCTTGGCTTGGCAGGTCAAGATCCAG 863
QY 222 TyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAla 241
Db 864 TACGACAAACAGCACGCTCTGTGGAGGAGCATCTCGACCCCACTGGTCTCTCAGGCA 923
QY 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
Db 924 GCCCACTGCTTCAGGAAACATACCGATGTGTTCAACTGGAAGGTGCGGGCAGGCTCAG 983
QY 262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsnPro 281
Db 984 AAACCTGGGCGAGCTTCCCATCCCTGGCTGTGGCCAAAGATCATCATTTGAATTCACCCC 1043
QY 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db 1044 ATGTACCCCAAGACAATGACATCGCCCTCATGAAGCTGCAGTCTCCACTCTCTTCTCA 1103
QY 302 GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro 321
Db 1104 GGCACAGTCAGGCCCATCTGTCTGCCCTTCTTTGATGAGGAGCTCACTCCAGCCACCCA 1163
QY 322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu 341
Db 1164 CTCTGGATCATTTGGATGGGCTTTACGAAGCAGAAATGGAGGGAAGATGTCTGACATACTG 1223
QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaTyrGln 361
Db 1224 CTGCAGGCGTCAGTCCAGGTCAATCCAGCACACCGGTCAATGCGAGACGATCGTACCAC 1283
QY 362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys 381
Db 1284 GGGGAAGTCAACCGAAGATGATGTGTGCGAGCATCCCGGAAGGGGGTGTGACACCTGC 1343
```


QY 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle 401
Db 1344 CAGGGTGACAGTGTGGGCCCTGATGTACCAATCTGACAGTGGCATGTGGTGGCAATC 1403
QY 402 ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSer 421
Db 1404 GTTAGCTGGGGCTATGGCTCGGGGGCCGAGCACCCAGGAGTATACCAAGGTCTCA 1463
QY 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db 1464 GCCTATCTCACTGGATCTACATGTCTGGAAGGCTGAGCTG 1505

RESULT 8

US-10-015-389A-274
; Sequence 274, Application US/10015389A
; Patent No. 6936436
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan I.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2830PIC48
; CURRENT APPLICATION NUMBER: US/10/015,389A
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 274
; LENGTH: 2063
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-015-389A-274

Alignment Scores:
Pred. No.: 1.67e-232 Length: 2063
Score: 2297.50 Matches: 429
Percent Similarity: 98.85% Conservative: 0
Best Local Similarity: 98.85% Mismatches: 0
Query Match: 98.10% Indels: 5
DB: 3 Gaps: 1

US-10-803-530-2 (1-435) x US-10-015-389A-274 (1-2063)

QY 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysProArg 21
Db 219 GATCCTGACAGTGATCAACTCTGAAAGAGCTGATGTCAAAACCCCTGGCCAAACCCCGT 278
QY 22 IleProMetGluThrPheArgLysValGlyIleProIleIleAlaLeuLeuSerLeu 41
Db 279 ATCCCCATGGAGACCTTCAGAAAGGTGGGATCCCCATCATCATAGCACTACTGACCTG 338
QY 42 AlaSerIleIleValValValValLeuIleLysValIleLeuAspLysTyrThrPheLeu 61
Db 339 GCGAGTATCATCTGGTGTGCTCATCAAGGTGATCTGATAAATACTACTTCTTC 398
QY 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys 81
Db 399 TGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 458
QY 82 ProLeuGluAspGluGluHisCysValLysSerPheProGluGlyProAlaValAla 101
Db 459 CCCTTGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 518

QY 102 ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp 121
Db 519 GTCCGCCCTCTCAAGGACCGATCCACACTGAGGGTGTGGACTCGGCACAGGGAACCTGG 578
QY 122 PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
Db 579 TTCTCTGCCCTGTTTCGACAACTTCACAGAAAGCTCTCGCTGAGACAGCGCTGTAGSCAGATG 638
QY 142 GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
Db 639 GGCTACAGC-----AGAGCTGTGAGATTGGCCAGACCAAGGATCTGGAT 683
QY 162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
Db 684 GTTGTGAAATTCAGAAAAACAGCAGGAGCTTCGCATGCGGAACCAAGTGGGCCCTGT 743
QY 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro 201
Db 744 CTCTCAGGCTCCCTGGTCTCCCTGCACCTGTCTGGATTCTTGGCTTGGCAGGTGACATCCAG 863
QY 202 ArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
Db 804 CGTGTGGTGGTGGGAGGAGGCTCTGTGGATTCTTGGCTTGGCAGGTGACATCCAG 863
QY 222 TyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAla 241
Db 864 TAGCAGAAAAACAGCACGCTCTGTGGAGGAGCATCTCGACCCCACTGGGTCTCTACGGCA 923
QY 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
Db 924 GCCCACTGTTTACGGAACATACCGATGTGTCACTGGAAGGTGGGAGGAGGCTCAGAC 983
QY 262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsnPro 281
Db 984 AAACCTGGGAGCTTCCCATCCCTGGCTGGTGGCCAGATCATCATATTGAATTCACACCC 1043
QY 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db 1044 ATGTACCCCAAGACAATGACATCGCCCTCATGAAGCTGCAATGCCACTCACTTCTCA 1103
QY 302 GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro 321
Db 1104 GGCACAGTCAGGCCCATCTGTCTGCCCTTCTTTGATGAGGAGCTCACTCCAGCCACCCCA 1163
QY 322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu 341
Db 1164 CTCTGGATCATTTGGATGGGGCTTTACGAAGCAGAGATGAGGAGGAGATGTCTGACATACTG 1223
QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaTyrGln 361
Db 1224 CTGAGGGGTGATCCAGTCAATGACAGCACAGGTGCAATGCAATGCAATGCAATGCAATG 1283
QY 362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyValValAspThrCys 381
Db 1284 GGGGAAGTCACCGAGAAGATGATGTGTGACAGGCATCCCGAAGGGGGTGTGGACACCTGC 1343
QY 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle 401
Db 1344 CAGGGTGAACAGTGGTGGGCCCTCATGTATACCAATCTGACCAAGTGGCATGTGGTGGCATC 1403
QY 402 ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSer 421
Db 1404 GTTAGCTGGGGCTATGGCTCGGGGGCCGAGCACCCAGGAGTATACCAAGGTCTCA 1463
QY 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db 1464 GCCTATCTCACTGGATCTACATGTCTGGAAGGCTGAGCTG 1505

RESULT 9

US-10-006-768A-274
; Sequence 274, Application US/10006768A
; Patent No. 6936697


```

; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2830PIC47
; CURRENT APPLICATION NUMBER: US/10/015,671A
; PRIOR FILING DATE: 2001-12-11
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 274
; LENGTH: 2063
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-015-671A-274

Alignment Scores:
Pred. No.: 1,67e-232 Length: 2063
Score: 2297.50 Matches: 429
Best Local Similarity: 98.85% Conservatives: 0
Query Match: 98.85% Mismatches: 0
DB: 98.10% Indels: 5
Gaps: 1

US-10-803-530-2 (1-435) x US-10-015-671A-274 (1-2063)

QY 2 AspProAspSerAspGlnProLeuAenSerLeuAspValLysProLeuArgLysProArg 21
Db 219 GATCCTGACAGTGATCAACCTCTGACACGCTCGATGTCAAAACCCCTCGCGAAACCCCGT 278

QY 22 IleProMetGluThrPheArgLysValGlyLeuProIleIleAlaLeuSerLeu 41
Db 279 ATCCCATGGAGACCTTCAGAAAGGTGGGATCCCATCATATAGCACTACTGAGCCTG 338

QY 42 AlaSerIleIleValValValLeuIleLysValIleLeuAspLysTyrPheLeu 61
Db 339 GCGAGTATCATATTGTGTCTCTCATCAAGGTGATCTGGATAAATACTACTTCCTC 398

QY 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys 81
Db 399 TCGCGGACGCTCTCACTTCATCCGAGGAGCAGCTGTGTGACGGAGAGCTGGACTGT 459

QY 82 ProLeuGlyGluAspGluHisCysValLysSerPheProGluGlyProAlaValAla 101
Db 459 CCCTTGGGGAGGACGAGGACACTGTGTCAAGAGCTTCCCGAAGGGCTTCAGTGGCA 518

QY 102 ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp 121
Db 519 GTCCGCTCTCCAGGACCGATCCACACTGCGAGGTGCTGGACTCGGCCACAGGAACTGG 578

QY 122 PheSerAlaCysPheAspAsnPhetheGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
Db 579 TTCTCTGCTGTTTCGACAACTTCACAGAACTCTCGCTGAGACAGCCTGTAGCGAGATG 638

QY 142 GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
Db 639 GGCTACAGC-----AGAGCTGTGAGATTGGCCAGACCCAGGATCTGGAT 693

QY 162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
Db 684 GTTGTGAAATTCAGAAAAACAGCAGGAGCTTCGCATGCGGAACCTCAAGTGGGCTGT 743

QY 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro 201
Db 744 CTCTCAGGCTCCCTTGGTCTCCCTGCACTGTCTTGGCTGTGGGAAGAGCCTGAAGACCCCC 803

QY 202 ArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
Db 804 CGTGTGTGGTGGGAGGAGGCTCTGTGATTTCTTGGCTTGGCAGGTACAGATCCAG 863

QY 222 TyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAla 241
Db 864 TACACAAAACAGCAGCTGTGAGGAGGAGCATCTGGACCCCACTGGGTCTCTACGGCA 923

QY 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
Db 924 GCCCACTGCTTCAGAAACATACCGATGTGTCACTGGAGGTGGCGGCGGCTCAGAC 983

; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2830PIC47
; CURRENT APPLICATION NUMBER: US/10/015,671A
; PRIOR FILING DATE: 2001-12-11
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 274
; LENGTH: 2063
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-015-393A-274

Alignment Scores:
Pred. No.: 1,67e-232 Length: 2063
Score: 2297.50 Matches: 429
Percent Similarity: 98.85% Conservatives: 0

US-10-015-393A-274

QY 262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsnPro 281
Db 984 AAACCTGGGACGCTTCCCATCCCTGGCTGTGGCCAAAGATCATCATCATTAATTCAACCC 1043

QY 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db 1044 ATGTACCCCAAGACAATGACATCGCCTCATGAAGCTGCAGTTCCTCCACTCCTTTCTCA 1103

QY 302 GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro 321
Db 1104 GGCACAGTCAGGCCCATCTGTCTGCCCTTTTGTATGAGGAGCTCATCCAGCCACCCCA 1163

QY 322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu 341
Db 1164 CTCTGGATCATATGGATGGGCTTTTACGAAGCAGAAATGGAGGGAAGATGTCTGACATACTG 1223

QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaTyrGln 361
Db 1224 CTGCAGGCGTCAGTCCAGGTCAITGACAGCACAGGTGCAATGCAGACGATGCGTACCAG 1283

QY 362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys 381
Db 1284 GGGGAAGTCAACCGAGAAGATGATGTGTGAGGCAATCCCGAAGGGGTGTGGACACTGC 1343

QY 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle 401
Db 1344 CAGGTGACAGTGGTGGGCCCCCTGATGTACCAATCTGACAGTGGCATGTGGTGGGCATC 1403

QY 402 ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSer 421
Db 1404 GTTAGCTGGGCTATGGCTCGGGGGCCCGAGCACCCAGGAGATATACACCAAGGTCTCA 1463

QY 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db 1464 GCCTATCTCACTGGATCTACAATGTCTGGAAGGCTGAGCTG 1505

RESULT 11
US-10-015-393A-274
; Sequence 274, Application US/10015393A
; Patent No. 6951737
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas P.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2830PIC46
; CURRENT APPLICATION NUMBER: US/10/015,393A
; CURRENT FILING DATE: 2002-06-10
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 274
; LENGTH: 2063
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-015-393A-274

Alignment Scores:
Pred. No.: 1,67e-232 Length: 2063
Score: 2297.50 Matches: 429
Percent Similarity: 98.85% Conservatives: 0

```

Best Local Similarity: 98.85% Mismatches: 0
Query Match: 98.10% Indels: 5
DB: 3 Gaps: 1

US-10-803-530-2 (1-435) x US-10-015-393A-274 (1-2063)

QY 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysProArg 21
DB 219 GATCTGACAGTATCAACCTCTGACAGCCTCGATGTCAACACCCCTGGCAACCCCGT 278

QY 22 IleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSerLeu 41
DB 279 ATCCCATGGAGACCTTCAGAAAGTGGGATCCCATCATCATAGCACTACTAGACCTG 338

QY 42 AlaSerIleIleIleValValLeuIleLysValIleLeuAspLysTyrTyrPheLeu 61
DB 339 GCGAGTATCATCTGTTGTTCTCATCAGGTGATTCTGGATAATACTACTTCTCTC 398

QY 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys 81
DB 399 TCGGGGAGCCTCTCCATCTCATCCGAGGAGCAGTGTGTGACGAGAGCTGACTGT 458

QY 82 ProLeuGlyGluAspGluHisCysValLysSerPheProGluGlyProIleValAla 101
DB 459 CCTTTGGGGAGGAGCAGAGCACTGTGTCAAGAGCTTCCCGAAGGGCCTGCACTGGCA 518

QY 102 ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp 121
DB 519 GTCCGCCCTCTCCAGGACGATCCACACTGCGAGGTGCTGGACTCGGCACAGGAACTGG 578

QY 122 PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
DB 579 TTCTCTGCTGTTTCGACAACTTCACAGAGCTCTCGTGAGACAGCCTGTAGGAGATG 638

QY 142 GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
DB 639 GGCTACAGC-----AGAGCTGTGGAGATTGGCCAGACCCAGGATCTGGAT 683

QY 162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
DB 684 GTTGTGTAATCACAGAAACAGCAGGAGCTTCGCATCGCGAACTCAAGTGGCCCTGT 743

QY 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro 201
DB 744 CTCTCAGGCTCCCTGGTCTCCCTGCACTGTCTGTGCTGTGGGAAGAGCCTGAAGACCCC 803

QY 202 ArgValValGlyGlyGluAlaSerValAspSerTrpTrpGlnValSerIleGln 221
DB 804 CGTGTGGTGGTGGGAGAGGCTCTGTGGATTCTTGGCCTTGGCAGGTGACGATCCAG 863

QY 222 TyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAla 241
DB 864 TAGCAACAAACAGCAGTCTGTGGAGGAGCATCTTGGACCCCACTGGTCTCTCAGGCA 923

QY 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
DB 924 GCCCACTGCTTCAGGAAACATACCATGATGTGTTCAACTGGAAGTGGGGCAGGCTCAGAC 983

QY 262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleIleGluPheAsnPro 281
DB 984 AAACCTGGGAGCTTCCCATCTCCCTGGCTGTGGCAAGATCATCATCAATTCAACCCC 1043

QY 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
DB 1044 ATGTACCCCAAGACAAATGACATCGCCCTCATGAAGCTGCAGTCCCACTCACTTCTCA 1103

QY 302 GlyThrValArgProIleCysLeuProPheAspGluGluLeuThrProAlaThrPro 321
DB 1104 GGCACAGTCAGGCCCACTCTGTGCTGCTTCTTGTGAGAGAGCTCACTCCAGCCACCCA 1163

QY 322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu 341
DB 1164 CTCTGGATCATTTGGATGGGCTTTACGAAGCAGATGGGAGGAGAGATGTCTGACATCTG 1223

QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGln 361
DB 1224 CTGCAGGCGTTCAGTCCAGTCAATTGACAGCACACCGTCAATGACAGCATGCGTACCAG 1283

QY 362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys 381
DB 1284 GGGGAAGTCCACGAGAAGATGATGTGACGATCCCGAAGGGGTGTGGACACCTGC 1343

QY 382 GlnGlyAspSerGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle 401
DB 1344 CAGGTGACAGTGGTGGCCCTGATGTACCAATCTGACAGTGGCATGTGTGGGCATC 1403

QY 402 ValSerTrpGlyTyrGlyCysGlyProSerThrProGlyValTyrThrLysValSer 421
DB 1404 GTTAGCTGGGCTATGGCTGGGGGCGCCGAGCACCCAGGAGTATACACCAAGTCTCA 1463

QY 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
DB 1464 GCCTATCTCAATGGATCTACAATGTCTGGAAGGCTGAGCTG 1505

RESULT 12

US-10-011-833A-274

; Sequence 274, Application US/10011833A

; Patent No. 6951920

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Botstein, David

; APPLICANT: Desnoyers, Luc

; APPLICANT: Eaton, Dan I.

; APPLICANT: Ferrara, Napoleone

; APPLICANT: Fong, Sherman

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, Christopher J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Hillan, Kenneth J.

; APPLICANT: Pan, James

; APPLICANT: Paoni, Nicholas F.

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; FILE REFERENCE: P2830P1C22

; CURRENT APPLICATION NUMBER: US/10/011,833A

; PRIOR APPLICATION DATE: 2002-06-25

; NUMBER OF SEQ ID NOS: 477

; SEQ ID NO 274

; LENGTH: 2063

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-011-833A-274

Alignment Scores:

Pred. No.: 1.67e-232 Length: 2063

Score: 2297.50 Matches: 429

Percent Similarity: 98.85% Conservativity: 0

Best Local Similarity: 98.85% Mismatches: 0

Query Match: 98.10% Indels: 5

DB: 3 Gaps: 1

US-10-803-530-2 (1-435) x US-10-011-833A-274 (1-2063)

QY 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysProArg 21
DB 219 GATCTGACAGTATCAACCTCTGACAGCCTCGATGTCAACACCCCTGGCAACCCCGT 278

QY 22 IleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSerLeu 41
DB 279 ATCCCATGGAGACCTTCAGAAAGTGGGATCCCATCATCATAGCACTACTAGACCTG 338

QY 42 AlaSerIleIleIleValValLeuIleLysValIleLeuAspLysTyrTyrPheLeu 61

```
Db 339 GCGAGTATCATATTGTTGTTCTCTCATCAAGGTGATTCTGGATAAATACTACTTCTC 398
Qy 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys 81
Db 399 TCGCGGAGCGCTCTCCACTTCATCCGAGGAAGCAGCTGTGTGACGGAGAGCTGGACTGT 458
Qy 82 ProLeuGlyGluAspGluGluHisCysValLysSerPheProGluGlyProAlaValAla 101
Db 459 CCCTTGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTCCCGGAAGGCCCTGCGAGTGGCA 518
Qy 102 ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp 121
Db 519 GTCCGCCCTCTCCAGGACCGATCCACACTGCAGGTGTGACTGGCCACAGGGAATGG 578
Qy 122 PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
Db 579 TTCTCTGCCTGTTTCGACAACTTCACAGAAGCTCTCGCTGAGACAGCCTGTAGCAGATG 638
Qy 142 GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
Db 639 GGCTACAGC-----AGAGCTGTGGAGATTGGCCCGACAGACCGAGATCTGGAT 683
Qy 162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
Db 684 GTTGTGAATCAGAGAAACAGCAGGAGCTTCGATGGGAATCAAGTGGGCCCTGT 743
Qy 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro 201
Db 744 CTCTCAGGCTCCCTGGTCTCCCTGCACCTGTCTTGCCTGTGGGAAGAGCCTGAAGACCCC 803
Qy 202 ArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
Db 804 CGTGTGTGGGTGGGAGGAGGCTCTGTGGATTCTTGGCCCTGGCAGGTGAGCATCCAG 863
Qy 222 TyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAla 241
Db 864 TACGACAAACAGCAGCTGTGTGGAGGAGGAGATCTGGACCCCACTGGGTCTCAGCGCA 923
Qy 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
Db 924 GCCCACTGCTTCAGGAACATACCGATGTGTCAACTGGGAAGGTGCGGCAGGCTCAGAC 983
Qy 262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleIleGluPheAsnPro 281
Db 984 AAACCTGGGCACTTCCCATCCCTGGCTGTGGCCAAAGATCATCATTTGAATTCAACCCC 1043
Qy 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db 1044 ATGTACCCCAAGACAATGACATCGCCCTCATGAAGCTGCAGTTCACCACTCTTCTCA 1103
Qy 302 GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro 321
Db 1104 GGCACAGTCAGGCCCACTGTCTGCCCTCTTTGATGAGGAGCTCACTCCAGCACACCCA 1163
Qy 322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyLysMetSerAspIleLeu 341
Db 1164 CTCTGGATCATTTGATGGGGCTTTTACGAAGCAGAATGGGGAAGATGTCTGCATACTG 1223
Qy 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGln 361
Db 1224 CTGCAGCGGTGAGTTCAGGTATTGACAGCACACGGTGCATATGACAGCATGCGTACAG 1283
Qy 362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys 381
Db 1284 GGGGAAGTCACCGAGAAGATGATGTGTGAGGCATCCCGNAGGGGTGTGGACACCTGC 1343
Qy 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle 401
Db 1344 CAGGGTGACAGTGTGTGGGCCCTTGATGTACCAATCTGACCAGTGGCATGTGGTGGGATC 1403
Qy 402 ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSer 421
Db 1404 GTTAGCTGGGGCTATGGCTGGGGGGGCCGAGNACCCAGGAGTATACCAAGGTCTCA 1463
```

```
Qy 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db 1464 GCCTATCTCAACTGGATCTACAATGTCTGGAAGGCTGAGCTG 1505
```

RESULT 13

```
US-10-006-041A-274
; Sequence 274, Application US/10006041A
; Patent No. 6951921
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan I.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2830P1C8
; CURRENT APPLICATION NUMBER: US/10/006,041A
; CURRENT FILING DATE: 2001-12-06
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 274
; LENGTH: 2063
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-006-041A-274
```

Alignment Scores:

```
Pred. No.: 1,67e-232 Length: 2063
Score: 2297.50 Matches: 429
Percent Similarity: 98.85% Conservative: 0
Best Local Similarity: 98.85% Mismatches: 0
Query Match: 98.10% Indels: 5
DB: 3 Gaps: 1
US-10-803-530-2 (1-435) x US-10-006-041A-274 (1-2063)
```

```
Qy 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysProArg 21
Db 219 GATCCTTGACAGTGATCAACCTCTGAACAGCCTCGATGCAACCCCTGCGCAACCCCGT 278
Qy 22 IleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSerLeu 41
Db 279 ATCCCATGAGAGACCTTCAGAAAGGTGGGATCCCCATCATATAGCACTACTGAGCTG 338
Qy 42 AlaserIleIleIleValValValLeuIleLysValIleLeuAspLysTyrTyrPheLeu 61
Db 339 GCGAGTATCATATTGTGGTTGCTCTCATCAAGGTGATTCTGGATAAATACTACTTCTC 398
Qy 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys 81
Db 399 TCGGGGAGCGCTCTCCACTTTCATCCGAGGAAGCAGCTGTGTGACGGAGAGCTGGACTGT 458
Qy 82 ProLeuGlyGluAspGluGluHisCysValLysSerPheProGluGlyProAlaValAla 101
Db 459 CCCTTGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTCCCCGAAGGCCCTGCGAGTGGCA 518
Qy 102 ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp 121
Db 519 GTCCGCCCTCTCCAGGACCGATCCACACTGCAGGTGTGACTCGGCCACAGGGAATGG 578
Qy 122 PheSerAlaCysPheAspAsnPheThrGluAlaIleuAlaGluThrAlaCysArgGlnMet 141
```

```
Db 579 TTCTCTGCTGTTTCGCAACTTCACAGAGCTCTCGCTGAGACAGCCTGTAGGAGATG 638
Qy 142 GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
Db 639 GGCTACAGC-----AGAGCTGTGGAGATTGGCCCCAGACGAGATCTGGAT 683
Qy 162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
Db 684 GTTGTGTAATACAGAAAAAGCAGAGAGCTTCGCATGCCGAATCAAGTGGGCCCTGT 743
Qy 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro 201
Db 744 CTCTCAGGCTCCCTGCTCTCCCTGCACGTCTTGGCTGTGGAGAGAGCTGAAGACCCC 803
Qy 202 ArgValValGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
Db 804 CGTGTGGTGGTGGGAGAGGCTCTCTGGATTCTTGGCTTGGCAGGTTCAGCATCCAG 863
Qy 222 TyrAspLysGlnHisValCysGlySerIleLeuAspProHisTrpValLeuThrAla 241
Db 864 TACGACAAACAGCAGCTCTGTGGAGGAGCATCTCTGGACCCCACTGGGTCTCTCAGGCA 923
Qy 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
Db 924 GCCCACTGCTTCAGAAAAATACCATCGATGTGTTCAACTGGAAGGTGCGGGCAGGCTCAGAC 983
Qy 262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsnPro 281
Db 984 AACTGGGAGCTTCCCATCCCTGGCTGTGGCAAGATCATCATTTGAATTCAAACCCC 1043
Qy 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db 1044 ATGTACCCCAAGACAATGACATGCCCTCATGAAGCTGCAGTTCCTCCACTCATTCTCA 1103
Qy 302 GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro 321
Db 1104 GGCAGAGTCAGGCCCATCTGTCTGCCCTTCTTTGATGAGGAGCTCACTCCAGCCACCCA 1163
Qy 322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu 341
Db 1164 CTCTGGATCATTGGATGGGGCTTTACGAAGCAGAAATGGAGGGAAGATGCTGACATCTG 1223
Qy 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaTyrGln 361
Db 1224 CTGAGGGCTGATGTCAGGTCAATTCAGACACACGTCGATGCATGCAGACGATGCGTACCAG 1283
Qy 362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys 381
Db 1284 GGGGAAGTCACCGAGAAGATGATGTGACAGGCATCCCGAAGGGGTGTGGACACCTGC 1343
Qy 382 GlnGlyAspSerGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle 401
Db 1344 CAGGGTGACAGTGTGGGCCCTGATGATACCAATCTGACCAAGTGCAGTGTGGTGGGCATC 1403
Qy 402 ValSerTrpGlyTyrGlyCysGlyProSerThrProGlyValTyrThrLysValSer 421
Db 1404 GTTACTGGGGCTATGGCTGGGGGGCCGAGCACCCAGGAGTATACACCAAGTCTCA 1463
Qy 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db 1464 GCCTATCTCAACTGGATCTACAAATGTCTGGAAGGCTGAGCTG 1505
```

RESULT 14

```
US-10-012-064A-274
; Sequence 274, Application US/10012064A
; Patent No. 6953841
```

GENERAL INFORMATION:

```
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan I.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
```

```
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2830P1C19
; CURRENT APPLICATION NUMBER: US/10/012,064A
; CURRENT FILING DATE: 2002-07-15
; PRIOR APPLICATION NUMBER: 60/098716
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098723
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098749
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098750
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098803
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/098821
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/098843
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/099536
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099596
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099598
; PRIOR FILING DATE: 1998-09-09
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 274
; LENGTH: 2063
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-012-064A-274
```

Alignment Scores:

Pred. No.:	1,67e-232	Length:	2063
Score:	2297.50	Matches:	429
Percent Similarity:	98.85%	Conservative:	0
Best Local Similarity:	98.85%	Mismatches:	0
Query Match:	98.10%	Indels:	5
DB:	3	Gaps:	1

US-10-803-530-2 (1-435) x US-10-012-064A-274 (1-2063)

```
Qy 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysProArg 21
Db 219 GATCTGACAGTGTCAACCTCTGAACAGCCTCGATGTCAAACCTCTGCGCAACCCCGT 278
Qy 22 IleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSerLeu 41
Db 279 ATCCCATGAGAGACCTTCAGAAAGTGGGGATCCCATCATCATAGCACTACTGAGCCCTG 338
Qy 42 AlaSerIleIleIleValValValIleValIleValIleValIleValIleValIleVal 61
Db 339 GCGAGTATCATATTGTGGTTGTCTCATCAAGGGTATTCTGGATAAATACTACTTCTTC 398
Qy 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys 81
Db 399 TGGGGGAGCTCTCCACTTTCATCCGAGGAGAGAGCTGTGTGACGAGAGAGCTGACTGT 458
Qy 82 ProLeuGlyGluAspGluGluHisCysValLysSerPheProGluGlyProAlaValAla 101
Db 459 CCGTTGGGGAGGACGAGGAGCAGTGTGTCAAGAGCTTCCCCGAAGGGCTCGAGTGGCA 518
Qy 102 ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp 121
```

```
Db 519 GTCCGGCTCTCCAGGACCGATCCACATGCGAGGTGCTGGACTCGGCCACAGGGAACCTGG 578
QY 122 PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
Db 579 TTCTCTGCTGTTTCAGCAACTTCACAGAGCTCTCGCTGACAGAGCTGTAGGAGATG 638
QY 142 GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
Db 639 GGCTACAGC-----AGAGCTGTGGAGATTGGCCACAGACCAGGATCTGGAT 683
QY 162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
Db 684 GTTGTGTAANTCACAGAAACAGCCAGAGCTTCGCATGCGGAACTCAAGTGGGCCCTGT 743
QY 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro 201
Db 744 CTCTCAGGCTCCCTGGTCTCCCTGACATGCTTGGCTGTGGAGAGAGCTGAAACCCCC 803
QY 202 ArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
Db 804 CGTGTGTGTGGTGGGAGAGAGGCTCTCTGTGATTCTTGGCTTGGCAGGTCCAGATCCAG 863
QY 222 TyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAla 241
Db 864 TACGACAAACAGCAGCTGTGTGGAGGGAGCATCTGGACCCCCACTGGGTCTCTCACGGCA 923
QY 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
Db 924 GCCCACTGCTTCAGGAACATACCGATGTGTTCACCTGGAAAGTGGGGAGGCTCAGAC 983
QY 262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsnPro 281
Db 984 AAATGGGCGAGCTTCCCATCCCTGGCTGTGGCCAAAGATCATCATCATTTCAACCCC 1043
QY 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db 1044 ATGTACCCCAAGAACAAATGACATCGCCCTCATGAAGCTGCAGTTCCCACTCACTTTCTCA 1103
QY 302 GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro 321
Db 1104 GGCAAGTCAAGGCCCATCTGTCTGCCCTTCTTTGATGAGGAGCTCACTCCAGCCACCCCA 1163
QY 322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu 341
Db 1164 CTCTGGATCATTTGGATGGGGCTTTTACGAAGCAGAAATGGAGGAGATGCTTGACATCTG 1223
QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaTyrGln 361
Db 1224 CTGCAGCGCTCAGTCCAGGTCATTGACAGACACCGGTGCAATGCAGACGATGGGTACCAG 1283
QY 362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys 381
Db 1284 GGGGAAGTCACCGAGAAGATGATGTGTGAGGCAATCCCGAAGGGGGTGTGGACACTGC 1343
QY 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle 401
Db 1344 CAGGGTGACAGTGTGGGCCCTTGATGTACCAATCTGACCACTGGCATGTGGTGGGCATC 1403
QY 402 ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSer 421
Db 1404 GTTAGCTGGGGCTATGGCTCGGGGGCCCGAGACCCCGAGGATATACACCAAGGTCTCA 1463
QY 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db 1464 GCCTATCTCAACTGGATCTACAATGTCTGGAAAGGCTGAGCTG 1505
```

RESULT 15

US-09-607-745-8

; Sequence 8, Application US/09607745

; Patent No. 6750034

; GENERAL INFORMATION:

; APPLICANT: Darrow, Andrew L

; APPLICANT: Qi, Jain-shen

```
; APPLICANT: Andrade-Gordon, Patricia
; TITLE OF INVENTION: DNA encoding human serine protease D-G
; FILE REFERENCE: ORT-1273
; CURRENT APPLICATION NUMBER: US/09/607,745
; CURRENT FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 1189
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Fusion gene
US-09-607-745-8
```

Alignment Scores:

Pred. No.:	3,72e-123	Length:	1189
Score:	1259.00	Matches:	227
Percent Similarity:	97.86%	Conservative:	2
Best Local Similarity:	97.01%	Mismatches:	5
Query Match:	53.76%	Indels:	0
DB:	3	Gaps:	0

US-10-803-530-2 (1-435) x US-09-607-745-8 (1-1189)

```
QY 202 ArgValValGlyGlyGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
Db 163 AAGATCGTTGGGGGCTATGCTAGATGTGGATTCTTGGCTTGGCAGGTCCAGATCCAG 222
QY 222 TyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAla 241
Db 223 TAGCAAAACAGCAGCATCTGTGGAGGGAGCATCTCTGGACCCCCACTGGGTCTCTCACGGCA 282
QY 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
Db 283 GCCCACTGCTTCAGGAACATACCGATGTGTTCACCTGGAAAGTGGGGAGGCTCAGAC 342
QY 262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsnPro 281
Db 343 AAATCGGGCAGCTTCCCATCCCTGGCTGTGGCCAAAGATCATCATCATTTCAACCCC 402
QY 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db 403 ATGTACCCCAAGAACAAATGACATCGCCCTCATGAAGCTGCAGTTCCCACTCACTTTCTCA 462
QY 302 GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro 321
Db 463 GGCAAGTCAAGGCCCATCTGTCTGCCCTTCTTTGATGAGGAGCTCACTCCAGCCACCCCA 522
QY 322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu 341
Db 523 CTCTGGATCATTTGGATGGGGCTTTTACGAAGCAGAAATGGAGGAGATGCTTGACATCTG 582
QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaTyrGln 361
Db 583 CTGCAGGCGTCAGTCCAGGTCAATTGACAGCACCGGTGCAATGCAGACGATGGGTACCCTG 642
QY 362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys 381
Db 643 GGGGAAGTCACCGAGAAGATGATGTGTGACAGCATCCCGAAGGGGGTGTGGACACCTGC 702
QY 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle 401
Db 703 CAGGGTGACAGTGTGGGCCCTTGATGTATACCAATCTGACCACTGGCATGTGGTGGGCATC 762
QY 402 ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSer 421
Db 763 GTTAGCTGGGGCTATGGCTCGGGGGCCCGAGACCCCGAGGGGTATACACCAAGGTCTCA 822
QY 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db 823 GCCTATCTCAACTGGATCTACAATGTCTGGAAAGGCTGAGCTG 864
```

Search completed: December 6, 2005, 23:11:01
Job time : 259 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: December 6, 2005, 19:52:16 ; Search time 4188 Seconds
(without alignments)
4859.690 Million cell updates/sec

Title: US-10-803-530-2
Perfect score: 2342
Sequence: 1 MDPDSQPLNSLDVKPLRPK.....VTKVSALYNIYVWKAEL 435

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues
Total number of hits satisfying chosen parameters: 82156650
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n model -DEV=xlp
-Q/cgn2_1/USPTO_spool_p/US10803530/runat 05122005 083216 2133/app query.fasta_1.583
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10803530 @CGN 1 1 8010 @runat 05122005 083216 2133 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST.*
1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_hic.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_est7.*
9: gb_gss1.*
10: gb_gss2.*
11: gb_gss3.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Query Match	Length	Description
1	1779	76.0	995	10 AY420489
2	1658.5	71.2	2177	4 AK078890
3	1427	60.9	995	10 AY420490
4	1424	60.8	995	10 AY420491
5	1233	52.6	844	2 BE531100
6	1203	51.4	895	5 BU522841
7	1158.5	49.5	975	2 BG288427

8	1108.5	47.3	920	2	BE516663
9	1054.5	45.0	617	3	BM795149
10	1042	44.5	956	5	BU523218
11	1037	44.3	602	7	CV570449
12	1023	43.7	582	3	BP253625
13	1015	43.3	591	7	CV028368
14	1009	43.1	573	3	BP260542
15	1006.5	43.0	583	3	BP253686
16	971	41.5	569	3	BM686292
17	951.5	40.6	804	2	BG222914
18	949.5	40.5	1007	2	BG386903
19	893	38.1	920	5	BQ22656
20	891	38.0	689	2	BG966811
21	890.5	38.0	879	2	BE516186
22	885	37.8	515	6	CB142902
23	881.5	37.6	645	2	BF118991
24	877	37.4	503	2	BE514663
25	875.5	37.4	787	2	BE515750
26	858	36.6	498	3	BM795128
27	856	36.5	548	6	CB159713
28	848	36.2	720	8	DN988024
29	848	36.2	744	3	BM008802
30	844	36.0	742	2	BG468475
31	839.5	35.8	665	7	CK950309
32	832.5	35.5	693	2	BG481239
33	825.5	35.2	935	2	BE515106
34	816	34.8	659	1	AI924527
35	812	34.7	827	7	CJ025679
36	808	34.5	648	1	AI924182
37	808	34.5	931	5	BU157248
38	797	34.0	563	3	BM773306
39	797	34.0	572	3	BM746729
40	797	34.0	578	3	BM747250
41	797	34.0	593	3	BM788452
42	797	34.0	596	3	BM772743
43	797	34.0	614	3	BM788163
44	797	34.0	633	3	BM764659
45	797	34.0	691	3	BM763697

ALIGNMENTS

RESULT 1
AY420489
LOCUS Homo sapiens TMPSR4 gene, VIRTUAL TRANSCRIPT, partial sequence, 995 bp DNA linear GSS 17-DEC-2003
DEFINITION genomic survey sequence.
ACCESSION AY420489
VERSION AY420489.1 GI:39776446
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 995)
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Direct Submision
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
PUBMED 14671302
REFERENCE 2 (bases 1 to 995)
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Direct Submision
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT This sequence as made by sequencing genomic exons and ordering them

based on alignment.

FEATURES		Location/Qualifiers	
source	1..995	/organism="Homo sapiens"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:9606"	
gene	<1..>995		
		/gene="TMPRSS4"	
		/locus_tag="HCM7241"	
ORIGIN			
Alignment Scores:			
Pred. No.:	1,15e-176	Length:	995
Score:	1779.00	Matches:	330
Percent Similarity:	99.70%	Conservative:	0
Best Local Similarity:	99.70%	Mismatches:	1
Query Match:	75.96%	Indels:	0
DB:	10	Gaps:	0
US-10-803-530-2 (1-435) x AY420489 (1-995)			
QY	52	LysValIleLeuAspLysTyrPheLeuCysGlyGlnProLeuHisPheIleProArg	71
Db	3	AAGGTGATTCTGGATAAATACTACTTCTCTCGGGCGACCTCTCCATTTATCCCGAGG	62
QY	72	LysGlnLeuCysAspGlyGluLeuAspCysProLeuGlyGluAspGluHisCysVal	91
Db	63	AAGCAGCTGTGTGACGGAGAGCTGGACTGTCTCTTGGGGAGCAGGAGCAGCTGTCTC	122
QY	92	LysSerPheProGluGlyProAlaValAlaValArgLeuSerLysAspArgSerThrLeu	111
Db	123	AGAGCTTCCCGAAGGGCTCGAGTGGAGTCCGCTCTCCAGAGCCGATCCACACTG	182
QY	112	GlnValLeuAspSerAlaThrGlyAsnTrpPheSerAlaCysPheAspAsnPhetrGlu	131
Db	183	CAGGTGCTGGACTCGGCCACAGGAACTGTTCTCTGCTGTTTCGACACTTTCACAGAA	242
QY	132	AlaLeuAlaGluThrAlaCysArgGlnMetGlyTyrSerSerLysProThrPheArgAla	151
Db	243	GCTCTCGCTGAGACAGCTGTAGGCAGATGGCTACAGCAGCAAAACCCACTTTCAGAGCT	302
QY	152	ValGluIleGlyProAspGlnAspLeuAspValValGluIleThrGluAsnSerGlnGlu	171
Db	303	GTGGAGATTGCCACAGACAGGATCTGGATGTGTGTTGAAATCACAGAAAACAGCCAGAG	362
QY	172	LeuArgMetArgAsnSerSerGlyProCysLeuSerGlySerLeuValSerLeuHisCys	191
Db	363	CTTCGATCGCGAACTCAAGTGGGCCCTGTCTCTCAGGCTCCCTGGTCTCCCTGCATGT	422
QY	192	LeuAlaCysGlyLysSerLeuLysThrProArgValValGlyGlyGluGluAlaSerVal	211
Db	423	CTTGCTGTGGGAAGAGCTGAAGACCCCGTGTGTGTGGTGTGGAGGCGCTCTGTG	482
QY	212	AspSerTrpProTrpGlnValSerIleGlnTyrAspLysGlnHisValCysGlyGlySer	231
Db	483	GATCTTGCCCTTGGCAGTCCAGATCCAGTACAGCAAAACAGCAGCTGTGGAGGAGG	542
QY	232	IleLeuAspProHisTrpValLeuThrAlaAlaHisCysPheArgLysHisThrAspVal	251
Db	543	ATCTCGACCCCACTTGGGTCTCTCAGCGACCCCACTGTCTCAGAAAACATCCGATGTG	602
QY	252	PheAsnTrpLysValArgAlaGlySerAspLysLeuGlySerPheProSerLeuAlaVal	271
Db	603	TTCACTGGAGGTGCGGGAGGCTCAGACAAATCGGGCAGCTTCCCATCCCTGGCTGTG	662
QY	272	AlaLysIleIleIleIleGluPheAsnProMetTyrProLysAspAsnAspIleAlaLeu	291
Db	663	GCCAAAGATCATCATATTGAATTCACCCCATGTATCCCCAAGACATGATCGCCCTC	722
QY	292	MetLysLeuGlnPheProLeuThrPheSerGlyThrValArgProIleCysLeuProPhe	311
Db	723	ATGAAGCTGCAGTTTCCCACTCACTTCTTCAGGGCAGTCAAGGCCCATCTGTCTGCCCTTC	782
QY	312	PheAspGluGluLeuThrProAlaThrProLeuTrpIleIleGlyTrpGlyPheThrLys	331
Db	783	TTTGATGAGGAGCTCACTCCAGCCACCCCACTCTGGATCATTTGGATGGGCTTTACGAAG	842
QY	332	GlnAsnGlyGlyLysMetSerAspIleLeuLeuGlnAlaSerValGlnValIleAspSer	351
Db	843	CAGAAATGGAGGGAAGATGTCTGCACATACTCTGTCAGGGCTCAGTCCAGGTCAATGACAGC	902
QY	352	ThrArgCysAsnAlaAspAspAlaTyrGlnGlyGluValThrGluLysMetMetCysAla	371
Db	903	ACACGGTGCATGTCAGACGATGCGTACCGGGGAACTCCACCGAGAAGATGATGTGTGCA	962
QY	372	GlyIleProGluGlyGlyValAspThrCysGln	382
Db	963	GGCATCCCGAAGGGGTGTGGACACCTGCCAG	995
RESULT 2			
AK078890			
LOCUS			
DEFINITION			
Mus musculus adult male colon cDNA, RIKEN full-length enriched library, clone:903062G02 product:SIMILAR TO TRANSMEMBRANE PROTEASE, SERINE 4 homolog [Mus musculus], full insert sequence.			
AK078890			
VERSION			
AK078890.1 GI:26098158			
KEYWORDS			
HTC; CAP trapper.			
SOURCE			
Mus musculus (house mouse)			
ORGANISM			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
REFERENCE			
1 Carninci, P. and Havaehizaki, Y.			
High-efficiency full-length cDNA cloning			
Meth. Enzymol. 303, 19-44 (1999)			
10349636			
REFERENCE			
2			
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.			
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes			
Genome Res. 10 (10), 1617-1630 (2000)			
11042159			
REFERENCE			
3			
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.			
RIKEN integrated sequence analysis (RISA) system-384-format sequencing Pipeline with 384 multicapillary sequencer			
Genome Res. 10 (11), 1757-1771 (2000)			
11076861			
REFERENCE			
4			
The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium.			
Functional annotation of a full-length mouse cDNA collection			
Nature 409, 695-690 (2001)			
5			
The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.			
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs			
Nature 420, 563-573 (2002)			
6 (bases 1 to 2177)			
REFERENCE			
AUTHORS			
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,			

Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

TITLE

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.c.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.jp/
URL: http://phantom.gsc.riken.jp/.

FEATURES

source

```

1..2177
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:9030622G02"
/db_xref="taxon:10090"
/clone="9030622G02"
/sex="male"
/tissue_type="colon"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
304..1507
/note="SIMILAR TO TRANSMEMBRANE PROTEASE, SERINE 4 homolog [Mus musculus] (SPR|AAH21368, evidence: FASTY, 92%ID, 100%length, match=1201) putative"
2154..2159
/note="putative"
2177
/note="putative"

```

misc_feature

polyA_signal

polyA_site

ORIGIN

Alignment Scores:

```

Pred. No.:      1,74e-164      Length:      2177
Score:          1668.50      Matches:      310
Percent Similarity: 80.05%      Conservative: 39
Best Local Similarity: 71.10%      Mismatches:  51
Query Match:      71.24%      Indels:      36
DB:              4              Gaps:         1

```

US-10-803-530-2 (1-435) x AK078890 (1-2177)

```

QY      1 MetAspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysPro 20
Db      304 ATGAGTCACAGAGTGGCCAAACCCCTCAACACCGGTATATTGCTCCCTTCGCAAAACC 363
QY      21 ArgIleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSer 40
Db      364 CGAAGGCCCCAGGAGACCTTCAAA----- 387
QY      41 LeuAlaSerIleIleIleValValValLeuIleLysValIleLeuAspLysTyrTyrPhe 60
Db      387 ----- 387
QY      61 LeuCysGlyGln-ProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAs 80
Db      388 -----AAGTCCCTTGACCTTCATTGAGAGGGCCAGTTGTTGACGGCCACCTTGA 438
QY      80 pCysProLeuGlyGluAspGluHisCysValLysSerPheProGluGlyProAlaVal 100
Db      439 CTGGCGCTCAGGGGAGGATGAGGAACACTGTGTCAAGGACTTCCTCGTGAAGGCCCGAGT 498

```

RESULT 3

AY420490 995 bp DNA linear GSS 17-DEC-2003
LOCUS Pan troglodytes TMRPS4 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION

```

QY      100 lAlaValArgLeuSerLysAspArgSerThrIleuGlnValLeuAspSerAlaThrGlyAs 120
Db      499 GGCAGTCCGGTCTCCCAAGGACAGATCCACCTCGAGTGTGGATGAGCCACAGGAC 558
QY      120 nTriPheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgG1 140
Db      559 CTGGCCCTCAGTCTGTTTCGACAACTTCACAGAGCACTGGCCAAAGACAGCCTGCAGACA 618
QY      140 nMetGlyTyrSerSerLysProThrPheArgAlaValGluLeuGlyProAspGlnAspLe 160
Db      619 GATGGCTATCAGACGCCAGCCCGCTTTTCAGAGCAGTGGAGATCCGTCAGATCAGAACC 678
QY      160 uAspValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyPr 180
Db      679 CCCTGTTGCTCAAGTCACAGAAACAGCCAGAACTTCAGGTGCAGAAATGGAAAGCAGATC 738
QY      180 cCysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysTh 200
Db      739 CTGCTCTCAGGCTCCCTGGTTTCCTTGGCTGCCCTTGACTGTGGAAGAGCCTGAAGAC 798
QY      200 rProArgValValGlyGlyGluAlaSerValAspSerTrpProTrpGlnValSerIl 220
Db      799 TCCTCGTGTGGTGGGTGGAGGCCCTCTGGATTCTTGGCGGTGGCAGTCCAGCAT 858
QY      220 eGlnTyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuTh 240
Db      859 CCAGTACAAACAGCAGCATGTCTGTGGTGGGAGCATCTCGATGTGCAAGAGCCTGATCT 918
QY      240 rAlaAlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySe 260
Db      919 AGCAGCCCACTGCTTCAGGAGTATCTTGATGTGTCAAGCTGGAAGGTTCAGGCGAGCTC 978
QY      260 rAspLysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAs 280
Db      979 AAACATACCTGGTAACTCTCCATCTTGCCTGTGGCCAAAGATCTTCACTCGCTGAACCC 1038
QY      280 nProMetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPh 300
Db      1039 TCCTCTGTACCCCAAGAGAGAGACATTCGCCCTTGTAAAGCTGCAGATCCCATCTCAAT 1098
QY      300 eSerGlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaTh 320
Db      1099 CTCAGCTCAGTCAGGCCCATCTGCTGCCCTTCTCTGATGAGGTCTTGTCCAGCCAC 1158
QY      320 rProLeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAsp1l 340
Db      1159 ACCAGTCTGGTTCATTTGGATGGGCTTTACAGAGAAACCGGAGGAAAGATGTTCTGACAT 1218
QY      340 eLeuLeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTy 360
Db      1219 GCTACTGCAGGCATCAGTCCAGTCAATTCAGACACACACCGTGCATATGACAGAGATGCCTA 1278
QY      360 rGlnGlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspTh 380
Db      1279 CGAAGGGGAAGTGACCGCTGAGATGCTGTGTCAGGTACCCACAGGGTGGCAGGACAC 1338
QY      380 rCysGlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValG1 400
Db      1339 CTGCCAGGTGACAGTGTGGGCTTTGATGTACCATTTCTGACAAGTGGCAGGTAGTAGG 1398
QY      400 yIleValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysVa 420
Db      1399 CATCGTAGCTGGGGCCATGATGTCGGCGGCCCAAGTACTCTCTGGAGTGTATACCAAGGT 1458
QY      420 lSerAlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db      1459 CACTGCTATCTCACTGGATCTACAATGTTTCGGAAGTCTGAGATG 1504

```

```

genomic survey sequence.
AY420490 AY420490.1 GI:39776447
GSS.
Pan troglodytes (chimpanzee)
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Pan.
1 (bases 1 to 995)
Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 995)
Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence is made by sequencing genomic exons and ordering them
based on alignment.
Location/Qualifiers
1..995
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>995
/feature="TMPRSS4"
/locus_tag="HCM7241"

gene
ORIGIN
Alignment Scores:
Pred. No.: 1.7e-139 Length: 995
Score: 1427.00 Matches: 273
Percent Similarity: 82.78% Conservative: 1
Best Local Similarity: 82.48% Mismatches: 57
Query Match: 60.93% Indels: 0
DB: 10 Gaps: 0

US-10-803-530-2 (1-435) x AY420490 (1-995)
Qy 52 LysValIleLeuAspLysTyTrPheLeuCysGlyGlnProLeuHisPheIleProArg 71
Db 3 AAGGTGATCTGGATAAATACTACTTCTCGCGGAGCCCTCCACCTTCATCCCGAGG 62
Qy 72 LysGlnLeuCysAspGlyGluLeuAspCysProLeuGlyGluAspGluHisCysVal 91
Db 63 AAGCAGCTGTGTGACGGAGCTGGACTGTCTCCCTGGGGAGGACGAGGACTGTGTC 122
Qy 92 LysSerPheProGluGlyProAlaValAlaValArgLeuSerLysAspArgSerThrLeu 111
Db 123 AAGAGCTTCCCGAAGGCCCTGGAGTGGCAGNNNNNNNNNNNNNNNNNNNNNNNN 182
Qy 112 GlnValLeuAspSerAlaThrGlyAsnTrpPheSerAlaCysPheAspAsnPheThrGlu 131
Db 183 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 242
Qy 132 AlaLeuAlaGluThrAlaCysArgGlnMetGlyTyTrSerSerLysProThrPheArgAla 151
Db 243 NNTCTCGCTGAGACAGCTGTAGCAGATGGGTACAGCAGCAAAACCCACTTTCAGAGCT 302
Qy 152 ValGluIleGlyProAspGlnAspLeuAspValValGluIleThrGluAsnSerGlnGlu 171
Db 303 GTGAGATTGGCCAGACCGAGGATCTGGATGTTGTTGAAATCACAGAAACAGCCAGGAG 362
Qy 172 LeuArgMetArgAsnSerSerGlyProCysLeuSerGlySerLeuValSerLeuHisCys 191

```

```

Db 363 CTTACATGCGGAACCTAAGTGGCCCTGTCTCTCAGGCTCCCTGGTCTCCTGCACGTG 422
Qy 192 LeuAlaCysGlyLysSerLeuLysThrProArgValValGlyGlyGluGluAlaSerVal 211
Db 423 CTTGCTGTGGGAGAGCGCTGANNNCCCGCTGTGTGGCGGGAGGAGGCTCTGTGTG 482
Qy 212 AspSerTrpProTrpGlnValSerIleGlnTyTrAspLysGlnHisValCysGlyGlySer 231
Db 483 GATTCCTTGGCCTTGGCAGGTCAGNATCCAGTACGNNNNNNNANNANGTCTGTGGAGGAGC 542
Qy 232 IleLeuAspProHisTrpValLeuThrAlaAlaHisCysPheArgLysHisThrAspVal 251
Db 543 ATCTGTGACCCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNAACATACCAGATGTG 602
Qy 252 PheAsnTrpLysValArgAlaGlySerAspLysLeuGlySerPheProSerLeuAlaVal 271
Db 603 TTCAACTGGAAGGTGCGGGCGGCTCAGACAACTGGGCAGCTTCCCATCCTGGCTGTG 662
Qy 272 AlaLysIleIleIleIleGluPheAsnProMetTyTrProLysAspAsnAspIleAlaLeu 291
Db 663 GCCAAGATCATCATATTGAATTCACCCCATGTACCCCAAGACAAATGATGATGCCTC 722
Qy 292 MetLysLeuGlnPheProLeuThrPheSerGlyThrValArgProIleCysLeuProPhe 311
Db 723 ATGAAGCTGCAGTTCNNNNNNNNNNNNNNNNNNNNNGCAGCTCAGGCCCATCTGTCTGCCCTTC 782
Qy 312 PheAspGluIleLeuThrProIleThrProLeuTrpIleIleGlyTrpGlyPheThrLys 331
Db 783 TTTGATGAGGAGCTCATCTCCAGCCACCCACTCTGGATCATTTGGATGGGGCTTTACGAAG 842
Qy 332 GlnAsnGlyGlyLysMetSerAspIleLeuLeuGlnAlaSerValGlnValIleAspSer 351
Db 843 CAGATGGGAGGAGATGTCGATCTGCTCAGCGGTCAGTCAGGTCATTGACAGC 902
Qy 352 ThrArgCysAsnAlaAspAlaTyTrGlnGlyValThrGluLysMetMetCysAla 371
Db 903 ACACGGTGCAATGCAGACGATCGTACCAGGGGGAAGTACCAGAGAAGATGATGTGTGCA 962
Qy 372 GlyIleProGluGlyGlyValAspThrCysGln 382
Db 963 GGCATCCCGAAGGGGGTGTGGACACCTGCCAG 995

RESULT 4
AY420491
LOCUS
DEFINITION
Mus musculus TmpRSS4 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
AY420491
AY420491.1 GI:39776448
GSS.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Murioidea; Muridae; Murinae; Mus.
1 (bases 1 to 995)
Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 995)
Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA

```

COMMENT This sequence as made by sequencing genomic exons and ordering them based on alignment.

FEATURES

source
1..995
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>995
/gene="TMPRSS4"
/locus_tag="HCM7241"

gene

ORIGIN

Alignment Scores:
Pred. No.: 3.53e-139 Length: 995
Score: 1424.00 Matches: 260
Percent Similarity: 87.61% Conservative: 30
Best Local Similarity: 78.55% Mismatches: 41
Query Match: 60.80% Indels: 0
DB: 10 Gaps: 0

US-10-803-530-2 (1-435) x AV420491 (1-995)

Qy 52 LysValIleLeuAspLysTyrPheLeuCysGlyGlnProLeuHisPheIleProArg 71
Db 3 AAGGTGATTCTGGATAAATACTACTTCTGCGGCAGTCCCTTTCATTCAGAGG 62
Qy 72 LysGlnLeuCysAspGlyGluLeuAspCysProLeuGlyGluAspGluHisCysVal 91
Db 63 GGCCAGTTGTGTGACGGCCACCTTGACTGCGCTCAGGGGAGGATGAGAACACTGTGTC 122
Qy 92 LysSerPheProGluGlyProAlaValAlaValArgLeuSerLysAspArgSerThrLeu 111
Db 123 AAGGACTTCCCTGAAAGCCCGAGTGGCAGTCCGCTCTCAAGCAGACAGATCCACCTG 182
Qy 112 GlnValLeuAspSerAlaThrGlyAsnTrpPheSerAlaCysPheAspAsnPheThrGlu 131
Db 183 CAGGTGCTGGATGACGCCACACAGGGACCTGGGCTTCAGTCTGTTCCGACAACTTCACAAA 242
Qy 132 AlaLeuAlaGluThrAlaCysArgGlnMetGlyTyrSerSerLysProThrPheArgAla 151
Db 243 GCATCGGCACAGACAGCCCTGCAGACAGATGGGCTATGACAGCCACCCCTTTCAGAGCA 302
Qy 152 ValGluIleGlyProAspGlnAspLeuAspValValGluIleThrGluAsnSerGlnGlu 171
Db 303 GTGGAGATCCCTCAGATCAGAACTCCCTGTGTCTCAAGTCACAGGAACAGCCAGGAA 362
Qy 172 LeuArgMetArgAsnSerSerGlyProCysLeuSerGlySerLeuValSerLeuHisCys 191
Db 363 CTTTCAGGTGACAGATGGAAGCAGATCTCGCTCTCAGGCTCCCTGTTTCTTGCACCTGC 422
Qy 192 LeuAlaCysGlyLysSerLeuLysThrProArgValValGlyGlyGluAlaSerVal 211
Db 423 CTTGACTGTGGAAGAGCCTGAAGACTCTCTGTGTGGTGGGTGGGAGGCCCTGTG 482
Qy 212 AspSerTrpProTrpGlnValSerIleGlnTrpAspLysGlnHisValCysGlyGlySer 231
Db 483 GATCTTGTGGCGTGGCAGTCCAGATCCAGTACACAGCAGATGTCTGTGGTGGAGC 542
Qy 232 IleLeuAspProHisTrpValLeuThrAlaAlaHisCysPheArgLysHisThrAspVal 251
Db 543 ATCTTGATCCCCACCTGGATCCTTCACAGAGCCCACTGTCTCAGGAAGTATCTTGATGTG 602
Qy 252 PheAsnTrpLysValArgAlaGlySerAspLysLeuGlySerPheProSerLeuAlaVal 271
Db 603 TCAAGCTGGAAGTCAAGGCGAGGCTCAAAATACATACTGGTAACCTCTCCATCCTTGCCTGTG 662
Qy 272 AlaLysIleIleIleIleGluPheAsnProMetTyrProLysAspAsnAspIleAlaLeu 291
Db 663 GCCAAGATCTTCACTGCTGCAACCAATCTCTGTATCCCAAGAGAGGACATTCGCCCTT 722
Qy 292 MetLysLeuGlnPheProLeuThrPheSerGlyThrValArgProIleCysLeuProPhe 311
Db 723 GTTAAAGCTGCAGATGCACCTCACATTCTCAGGCTCAGTCAAGGCCCATCTGCCTGCCCTTC 782

Qy 312 PheAspGluLeuLeuThrProAlaThrProLeuTrpIleIleGlyTyrGlyPheThrLys 331
Db 783 TCTGATGAGGTGCTTGTCCAGGCACACACAGTCTGGGTCTATTGGATGGGCTTTTACAGAA 842
Qy 332 GlnAsnGlyGlyLysMetSerAspIleLeuLeuGlnAlaSerValGlnValIleAspSer 351
Db 843 GAAACGGAGGAAAGATGTCTGACATGCTACTGAGGCATCAGTCCAGGTTCATCGACAGC 902
Qy 352 ThrArgCysAsnAlaAspAspAlaTyrGlnGlyGluValThrGluLysMetMetCysAla 371
Db 903 ACACGGTCAATGCAGAGGATGCTTACGAAGGGGAGTGAACCGCTGAGATGCTGTGTGCA 962
Qy 372 GlyIleProGluGlyGlyValAspThrCysGln 382
Db 963 GGTACCCACACAGGGTGGCAAGGACACCTGCCAG 995
RESULT 5
BE531100
LOCUS 601278466F1 NIH_MGC_39 Homo sapiens cDNA clone IMAGE:3610337 5',
DEFINITION mRNA sequence.
ACCESSION BE531100
VERSION BE531100.1 GI:9759745
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
REFERENCE 1 (bases 1 to 844)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaops-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM266 row: d column: 18
High quality sequence stop: 712.
Location/Qualifiers
1..844
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3610337"
/tissue_type="adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_39"
/note="Organ: pancreas; Vector: pOTB7; Site: 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dt priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed
by Ling Hong in the laboratory of Gerald M. Rubin
(University of California, Berkeley) using ZAP-cDNA
synthesis kit (Stratagene) and Superscript II RT (Life
Technologies)."

ORIGIN
Alignment Scores:
Pred. No.: 4.07e-119 Length: 844
Score: 1233.00 Matches: 254
Percent Similarity: 91.49% Conservative: 4
Best Local Similarity: 90.07% Mismatches: 14
Query Match: 52.65% Indels: 15
DB: 2 Gaps: 2
US-10-803-530-2 (1-435) x BE531100 (1-844)

```
QY 40 SerLeuAlaSerIleIleIleValValValValLeuIleIleValIleLeuAspLysTyrTyr 59
Db 1 AGCTGGCGAGTATCATCTTGTGGTTGTCTCATCAAGTGATTCGTGATAAATACATAC 60
QY 60 PheLeuCysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeu 79
Db 61 TTCCTCTGGGGCAGGCTCTCCACTTCATCCCGAGGAACAGCTGTGTGACGAGAGCTG 120
QY 80 AspCysProLeuGlyGluAspGluGluHisCysValLysSerPheProGluGlyProAla 99
Db 121 GACTGTCCCTTGGGGGAGACGAGGAGCACTGTGTCAAGAGCTTCCCGAAGGGCTGCA 180
QY 100 ValAlaValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGly 119
Db 181 GTGGCAGTCCGCTCTCCAAAGACCGATCCACACTGCAGGTCTCGACTCGGCCACAGGG 240
QY 120 AsnTrpPheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArg 139
Db 241 AACTGGTCTCTGCCCTTTCGACAACTTCAGAAAGCTCTCGCTGAGACAGCCTGTAGG 300
QY 140 GlnMetGlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAsp 159
Db 301 CAGATGGGTACAGCAGCAAAACCCACTTTCAGAGCTGTGGAGATTGGCCCGACAGCAGAT 360
QY 160 LeuAspValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGly 179
Db 361 CTGGATGTGTGTAATCAGAAACAGCCAGGAGCTTCGCATCGCGAACTCAAGTGGG 420
QY 180 ProCysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLys 199
Db 421 CCTGTCTCTCAGGCTCCCTGCTCTCCCTGCACTGTCTGTGCTGTGGGAAGCCTGAAG 480
QY 200 ThrProArgValValGlyGlyGluGluAlaSerValAspSerTrpTrpGlnValSer 219
Db 481 ACCCCCGGTGTGGTGGGGAGGAGGCTCTGTGGATTCTTGGCCTTGGCAGTGCAGC 540
QY 220 IleGlnTrpAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeu 239
Db 541 ATCCAGTAGCAACAAACAGACGTCTGTGGAGGAGCATCTCGGACCCCGCATCTGAATCC 600
QY 240 ThrAlaAlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGly 259
Db 601 ACCGCAGC-CACTGTCTCAGGAAACATACCATGTGTCACTGGAAGGTGGGGCAGGC 659
QY 260 SerAspLysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPhe 279
Db 660 TCAGACAA-CTGGGCGAGCTTCCCATCCCTGGGGCTGTGGCCAGATCATCATTTGAATCC 718
QY 280 AsnProMetTyrProLysAspAsnAlaLeuMetLysLeuGlnPhePro----- 297
Db 719 AACCTGTGT---CCCCAGACAT-GACATCGCCCTCATGAAGCTGCG-TTCCCTCTCACTT 773
QY 298 -----LeuThrPheSerGlyThrValArgProIleCysLeuPro 310
Db 774 CTCGGCGCGCGCGGCTCTGTGTGCTTTGTGGGGCCCT-CGCCCCCTCTGTCTGGGG 832
QY 311 PhePhe 312
Db 833 GGGTTC 838

RESULT 6
BUS522841
LOCUS
DEFINITION AGENCOURT_10158166 NCI_CGAP_Co24 Mus musculus cDNA clone
IMAGE:6529422 5', mRNA sequence.
ACCESSION BUS522841
VERSION BUS522841.1 GI:22833279
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
```

```
REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM4129 row: i column: 06
High quality sequence stop: 665.
FEATURES
source
1..895
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:6529422"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP Co24"
/notes="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
```

ORIGIN

```
Alignment Scores:
Pred. No.: 6,54e-116 Length: 895
Score: 1203.00 Matches: 217
Percent Similarity: 89.17% Conservative: 30
Best Local Similarity: 78.34% Mismatches: 29
Query Match: 51.37% Indels: 1
DB: 5 Gaps: 0
US-10-803-530-2 (1-435) x BUS522841 (1-895)
QY 154 IleGlyProAspGlnAspLeuAspValValGluIleThrGluAsnSerGlnGluLeuArg 173
Db 23 GTCCGTCCAGATCAGAACCTCTGTGTCTCAAGTCACAGAAACAGCCAGGAACCTTCAG 82
QY 174 MetArgAsnSerSerGlyProCysLeuSerGlySerLeuValSerLeuHisCysLeuAla 193
Db 83 GTGCAGATGGAGCAGATCTCTCTCAGGCTCCCTGGTTTCTTGGCTGCGCTTGAC 142
QY 194 CysGlyLysSerLeuLysThrProArgValValGlyGlyGluGluAlaSerValAspSer 213
Db 143 TGTGGAAGAGAGCTGAAGACTCTCGTGTGTGGTGGGTGGGGTGGAGGCCCTGTGGATTCT 202
QY 214 TrpProTrpGlnValSerIleGlnTrpAspLysGlnHisValCysGlyGlySerLeu 233
Db 203 TGGCGGTGGCAGGTGAGATCCAGTACAGTACAAAGCAGCATGTCTGTGGTGGAGCATCCTG 262
QY 234 AspProHisTrpValLeuThrAlaAlaHisCysPheArgLysHisThrAspValPheAsn 253
Db 263 GATCCCACTGGATCCTCACAGCAGCCCACTGCTTCAGGAAGTATCTTGTATGTCAAGC 322
QY 254 TrpLysValArgAlaGlySerAspLysLeuGlySerPhePheProSerLeuAlaValAlaLys 273
Db 323 TGGAAAGGTGAGGCGAGGCTCAAAACATACTGGGTAACTCTCCATCTCTGCTGTGGCCAA 382
QY 274 IleIleIleIleGluPheAsnProMetTyrProLysAspAsnAspIleAlaLeuMetLys 293
Db 383 ATCTTCATCGCTGAACCAATCCTCTGACCCCAAGAGAGGACATTCGCCCTTTTAAG 442
QY 294 LeuGlnPheProLeuThrPheSerGlyThrValArgProIleCysLeuProPhePheAsp 313
Db 443 CTGCAGATGCCACTCACATTCTCAGGCTCAGTCAGGCCCATCTGCCTGCCCTTCTCTGAT 502
```

Qy	314	GluGluLeuThrProAlaThrProLeuTrpIleIleGlyTrpGlyPheThrLysGlnAsn	333		
Db	503	GAGGTGCTTGTCCAGCCACACAGCTTGGGTCTATGGATGGGCTTTACAGAGAAAC	562		
Qy	334	GlyGlyLysMetSerAspIleLeuLeuGlnAlaSerValGlnValIleAspSerThrArg	353		
Db	563	GGAGGAAGATGTTCTGACATCTACTGCGAGCATCAGTCCAGGTCATCGACAGCACGG	622		
Qy	354	CysAsnAlaAspAspAlaThrGlnGlyValThrGluLysMetMetCysAlaGlyIle	373		
Db	623	TGCAATGCAGAGGATGCTTACAGAGGGAAGTGCACGCTGAGATGCTGTGCGAGGTACC	682		
Qy	374	ProGluGlyGlyValAspThrCysGlnGlyAspSerGlyGlyProLeuMetTyrGlnSer	393		
Db	683	CCACAGGGTGCCAGGACACCTGCCAGGTCACAGTGGTGGGCTTTGATGATCACTTCT	742		
Qy	394	AspGlnTrpHisValValGlyIleValSerTrpGlyTyrGlyCysGlyGlyProSerThr	413		
Db	743	GACAAGTGGCAGGTAGTAGGCATCTGAGCTGGGGCCATGGATGGCGGGCCCAAGTACT	802		
Qy	414	ProGlyValTyrThrLysValSerAlaTyrLeuAsnTrpIleTyrAsn	429		
Db	803	CCCTGGAGTGATACCAAGGTCACTGCCTATCTCAACTGGATCTACAAT	851		
RESULT 7					
LOCUS	BG288427	60238091F1 NIH_MGC_93 Homo sapiens cDNA clone IMAGE:4516990 5',	EST 21-FEB-2001		
DEFINITION	mRNA sequence.				
ACCESSION	BG288427				
VERSION	BG288427.1	GI:13043459			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE					
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/.				
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)				
JOURNAL	Unpublished (1999)				
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM10409 row: e column: 23 High quality sequence stop: 659.				
FEATURES					
source	Location/Qualifiers				
	1..975				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/clone="IMAGE:4516990"				
	/tissue_type="transitional cell papilloma, cell line"				
	/lab_host="DH10B (phage-resistant)"				
	/clone_lib="NIH MGC 93"				
	/note="Organ: bladder; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.7 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."				
ORIGIN					
Alignment Scores:					
Pred. No.:	3.72e-111	Length:	975		
Score:	1158.50	Matches:	221		
Percent Similarity:	96.94%	Conservative:	1		
Best Local Similarity:	96.51%	Mismatches:	4		

Query Match:	49.47%	Indels:	4
DB:	2	Gaps:	1
US-10-803-530-2 (1-435) x BG288427 (1-975)			
Qy	196	LysSerLeuLysThrProArgValValGlyGlyGluAlaSerValAspSerTrpPro	215
Db	3	AAGAGCCTGAAGACCCCTGCTGGTGGTGGGAGGAGGCTCTGTGGATTCCTGGCCT	62
Qy	216	TrpGlnValSerIleGlnTyrAspLysGlnHisValCysGlyGlySerIleLeuAspPro	235
Db	63	TGGCAGGTTCAGCATCCAGTACGACAAACAGCACGCTCTGTGGAGGGAGCATCTCGACCCC	122
Qy	236	HisTrpValLeuThrAlaAlaHisCysPheArgLysHisThrAspValPheAsnTrpLys	255
Db	123	CACCTGGGTCTCAGCGACGCCCTGCTTCCAGGAACATACCGATGTGTTCAACTGGAAG	182
Qy	256	ValArgAlaGlySerAspLysLeuGlySerPheProSerLeuAlaValAlaLysIleIle	275
Db	183	GTGGGGGAGGCTCAGACAAACTGGGAGGCTTCCCATCCCTGGCTGTGGCCACAGATCATC	242
Qy	276	IleIleGluPheAsnProMetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGln	295
Db	243	ATCATTTGAATTCACCCCTGATGATCCCAAGACATGACATCGCCCTCATGAAGCTGCAG	302
Qy	296	PheProLeuThrPheSerGlyThrValArgProLysCysLeuProPhePheAspGluGlu	315
Db	303	TTCCCACTCACTTCTCAGGCACAGTCAGGCCCATCTGCTGCCCTTCTTTGATGAGGAG	362
Qy	316	LeuThrProAlaThrProLeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGly	335
Db	363	CTCACTCCAGCCACCCCACTCTGGATCATTTGGATGGGCTTTACGAAGCAGAATCGAGGG	422
Qy	336	LysMetSerAspIleLeuLeuGlnAlaSerValGlnValIleAspSerThrArgCysAsn	355
Db	423	AAGATGCTGACATCTGCTGAGCGCTGCTCCAGGCAATTCAGACGACACGCGTGCAT	482
Qy	356	AlaAspAspAlaTyrGlnGlyGluValThrGluLysMetMetCysAlaGlyIleProGlu	375
Db	483	CGACAGCATGCTGATCAGGGGAGTACCAGAGAGATGATGTGTGAGGCATCCCGAA	542
Qy	376	-GlyGlyValAspThrCysGlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspG	395
Db	543	GGGGGGTGTGGACACCTGCCA-GGTGACAGTGGTGGGCGCCCTGATGTACCAATCTGACCA	601
Qy	395	nTrpHisValValGlyIleValSerTrpGlyTyrGlyCysGlyGlyProSerThrProG	415
Db	602	GTGGCATGTGTGGGCATCGTTAGCTGGGGCTATGGCTGCGGGGCC---GAGCACCAGG	658
Qy	415	yValTyr-ThrLysValSerAla	422
Db	659	AGTTTACCACCAAGGTCTCAAGC	681
RESULT 8			
LOCUS	BE616663	920 bp	mRNA linear EST 24-AUG-2000
DEFINITION	601278696F1 NIH_MGC_39 Homo sapiens cDNA clone IMAGE:3610990 5', mRNA sequence.		
ACCESSION	BE616663		
VERSION	BE616663.1	GI:9898262	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE			
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/.		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: ATCC		

cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

Plate: L1CM267 row: o column: 23
 High quality sequence start: 5
 High quality sequence stop: 819.

FEATURES

Location/Qualifiers
 1..920
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3610990"
 /tissue_type="adenocarcinoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 39"
 /note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
 Site_2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Library constructed
 by Ling Hong in the laboratory of Gerald M. Rubin
 (University of California, Berkeley) using ZAP-cDNA
 synthesis kit (Stratagene) and Superscript II RT (Life
 Technologies)."

ORIGIN

Alignment Scores:
 Pred. No.: 6.49e-106 Length: 920
 Score: 1108.50 Matches: 255
 Percent Similarity: 86.21% Conservative: 9
 Best Local Similarity: 83.33% Mismatches: 27
 Query Match: 47.33% Indels: 20
 DB: 2 Gaps: 5

US-10-803-530-2 (1-435) x BE616663 (1-920)

QY 39 LeuSerLeuAla-SerIleIleValValValLeuIleLysValIleLeuAspLysTy 58
 Db 15 ATGAGCTGGCGGAGGTATCATTTGTTGCTCTCATCAAGGTATCTGGATAATA 74
 QY 58 rTyPheLeuCysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyG 78
 Db 75 CTACTTCTCTCGCGGCGAGCTCTCCACTTCATCCGAGGAGGAGCTGTGTGACGAGA 134
 QY 78 uLeuAspCysProLeuGlyGluAspGluGluHisCysValLysSerPheProGluGlyPr 98
 Db 135 GCTGAGCTGTCCTGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 194
 QY 98 oAlaValAlaValArgLeuSerLys-AspArgSerThrLeuGlnValLeuAspSerAla 118
 Db 195 TGCAGTGGCAGTCCGCTCTCAAGGGACCGATCCACATGTCAGGTCTGGACTCGGCCA 254
 QY 118 hrGlyAsnTrpSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaC 138
 Db 255 CAGGGAACCTGGTCTCTGCTCTGTTTCGACAACTTCACAGAGCTCTCGCTGACAGCCT 314
 QY 138 ySArgGlnMetGlyTyTrSerLysProThrPheArgAlaValGluIleGlyProAspG 158
 Db 315 GTAGGCAGATGGCTACAGCAGCAACCCACTTTCAGAGCTGTGGAGATTGGCCCGAC 374
 QY 158 lnAspLeuAspValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerS 178
 Db 375 AGGATCTGGATGTTGTTGAAATCACAGAAACAGCCAGGAGCTTCGATCGGAGACTCAA 434
 QY 178 erGlyProCysLeuSerClySerLeuValSerLeuHisCysLeuAlaCys-GlyLysSer 197
 Db 435 GTGGGCCCTGTCTCTCAGGCTCCCTGCTCTCCCTGCACTGTCTTGGCTGTGGGAAAGAG 494
 QY 198 LeuLysThrProArgValValGlyGlyGluAlaSerValAspSerTrpPro-TrpG 217
 Db 495 CTGAAGAACCCCGTGTGGTGGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 554

QY 217 nValSerIleGlnTyTr-AspLysGlnHisValCysGlySerIleLeuAspProHist 237
 Db 555 GGTGAGCATTCAGTACGACAAACAGCAGCGTCTGTGGAGGAGCATCTGGA-CCCACT 613
 QY 237 rpValLeuThrAlaAlaHisCysPheArgLysHisThrAspValPheAsn---TrpLysV 256
 Db 614 GGGTCTCTCAGGA-AGCCACTGTTTCCGAAACATACCGAATGTTGTTCACATTGGAAGG 672
 QY 256 al-ArgAlaGlySer-Asp---LysLeuGlySer-----PheProSerLeuAlaValAl 272
 Db 673 TGCCGGGCGAGCTCAAGAACAAAACTGGGCAAGGTTCCCAATACCTGCTGTGGGCCAA 732
 QY 272 aLysIleIleIleIleGluPheAsnProMetTyTrProLysAspAsn-AspIleAlaLeuM 292
 Db 733 AGATCCTTCATCATTTGAATTCACCCCATGTATCCCAAGACAAATTGACATGCCCTCA 792
 QY 292 etLysLeuGlnPheProLeuThrPheSerGlyThrValArgProIleCysLeuProPheP 312
 Db 793 TGAAGCTGAAGTT-CCACTCATTCTTCAGGCACAGCAGGCCATC-TGTTTGCT---C 847
 QY 312 heAspGluGluLeuThrProAlaThrProLeuTrpIleIleGlyTrpGlyPheThrLysG 332
 Db 848 TTGATGAGGAGTC-CATCCAGCACCCCATCTGTGTCATGTT-----GGGTTACAGCC 900
 QY 332 ln 332
 Db 901 AG 902

RESULT 9
 BM795149
 LOCUS
 DEFINITION
 K-EST0076707 S21SNU520 Homo sapiens cDNA clone S21SNU520-27-H07 5',
 mRNA sequence.
 ACCESSION
 VERSION
 BM795149.1 GI:19143381
 EST.
 KEYWORDS
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 617)
 Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
 Kim,I.S.
 21c Frontier Korean EST Project 2001
 Unpublished (2002)
 CONTACT: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.kribb.re.kr
 Plate: 27 row: H column: 07
 High quality sequence stop: 617.

FEATURES

Location/Qualifiers
 1..617
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="S21SNU520-27-H07"
 /sex="F"
 /tissue_type="Stomach"
 /cell_type="Floating aggregates"
 /cell_line="SNU-520"
 /lab_host="Top10P"
 /clone_lib="S21SNU520"
 /note="Organ: Stomach; Vector: pTZ18RP1; Site_1: EcoRI;
 Site_2: NotI; The poly (A) + RNA was dephosphorylated with
 bacterial alkaline phosphatase (BAP) and then decapped
 with tobacco acid pyrophosphatase (TAP). The decapped

intact mRNA was ligated with DNA-RNA linker including EcoR I site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transposition of competent cells E. coli Top10F⁺ by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

ORIGIN

Alignment Scores:
Pred. No.: 1,798-100 Length: 617
Score: 1054.50 Matches: 199
Percent Similarity: 97.56% Conservative: 1
Best Local Similarity: 97.07% Mismatches: 4
Query Match: 45.03% Indels: 1
DB: 3 Gaps: 1

US-10-803-530-2 (1-435) x BM795149 (1-617)

QY	76	AspGlyGluLeuAspCysProLeuGlyGluAspGluHisCysVallySerPhePro	95
DB	2	GACGAGAGCTGGATGTCCTTGGGGAGGACGAGGACATGTCAGAGCTTCCCC	61
QY	96	GluGlyProAlaValAlaValArgLeuSerLysAspArgSerThrLeuGlnValLeuAsp	115
DB	62	GAAGGGCTGAGTGGCAGTCGCTCTCCAGGACCGATCCACACTGGAGTGTGGAC	121
QY	116	SerAlaThrGlyAsnTrpPheSerAlaCysPheAspAsnPhetrGluAlaLeuAlaGlu	135
DB	122	TCGGCCACAGGAACCTGCTCTGCTCTTCGCACTTCACAGAACTTCGCTGAG	181
QY	136	ThrAlaCysArgGlnMetGlyTyrSerSerLysProThrPheArgAlaValGluLeuGly	155
DB	182	ACAGCTGTAGGACATGGCTACAGCAGCAAAACCCACTTTCAGAGCTGTGGAGATTGGC	241
QY	156	ProAspGlnAspLeuAspValValGluLeuThrGluAsnSerGlnGluLeuArgMetArg	175
DB	242	CCAGACCGAGATCTGGATGTTGTTGAATACAGAAACAGCAGGAGCTTCGATCGCG	301
QY	176	AnSerSerGlyProCysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGly	195
DB	302	AACTCAAGTGGCCCTGCTCTCAGGCTCCCTGGTCTCCCTGCACCTCTTGCCTGTGG	361
QY	196	LysSerLeuLysThrProArgValValGly---GlyGluGluAlaSerValAspSerTrp	214
DB	362	AAAGAGCTGAAGACCCCGTGTGTGTGGTGGGGAAGGANGCCCTCTGTGGATTCTTGG	421
QY	215	ProTrpGlnValSerileGlnTyrAspLysGlnHisValCysGlyGlySerileLeuAsp	234
DB	422	CCTTGGCAGGTGAGCAGATACAGTACGACAAACAGCAGCTCTGTGGAGGAGCATCTGGAC	481
QY	235	ProHisTrpValLeuThrAlaAlaHisCysPheArgLysHisThrAspValPheAsnTrp	254
DB	482	CCCCATGGGTCTCAGCGAGCCCACTGCTTCAGGAACATACCCGATGTGTTCACTGG	541
QY	255	LysValArgAlaGlySerAspLysLeuGlySerPheProSerLeuAlaValAlaLysIle	274
DB	542	AAGGTGGGCGAGCTCAAAACAACTGGGCGAGCTTCCCATCCCTGGCTGTGGCCAAGATC	601
QY	275	IleIleIleGluPhe	279
DB	602	ATCATCATTTGAATTC	616

RESULT 10

BUS23218
LOCUS BUS23218 956 bp mRNA linear EST 13-SEP-2002
DEFINITION AGENCOURT_10154487 NCI_CGAP_C024 Mus musculus cDNA clone
IMAGE:6529864 5', mRNA sequence.

ACCESSION
VERSION BUS23218.1
KEYWORDS GI:22833656
SOURCE EST.
ORGANISM Mus musculus (house mouse)

Mus musculus (house mouse)

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.

1 (bases 1 to 956)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs@mail.nih.gov

Tissue Procurement: The Cepko Laboratory

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM14130 row: k column: 16

High quality sequence stop: 706.

Location/Qualifiers

FEATURES
source

1..956
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:6529864"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP C024"
/note="Organ: colon; Vector: pCMV-SPORT6; Site: 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI CGAP Library."

ORIGIN

Alignment Scores:

Pred. No.: 7,238-99 Length: 956
Score: 1042.00 Matches: 213
Percent Similarity: 81.25% Conservative: 34
Best Local Similarity: 70.07% Mismatches: 48
Query Match: 44.49% Indels: 9
DB: 5 Gaps: 4

US-10-803-530-2 (1-435) x BUS23218 (1-956)

QY	96	GluGlyProAlaValAlaValArgLeuSerLysAspArgSerThrleuGlnValleuAsp	115
DB	46	GAAGAGCCGGAGTGGCAGTCCGGCTCCCAAGGACAGATCCACCTGCAGGTGCTGGAT	105
QY	116	SerAlaThrGlyAsnTrpPheSerAlaCysPheAspAsnPhetrGluAlaLeuAlaGlu	135
DB	106	GGAGCCACAGGACCTGGGCCTCAGTCGTTCGACAACTTCACAGAGCACTGGCCAAG	165
QY	136	ThrAlaCysArgGlnMetGlyTyrSerSerLysProThrPheArgAlaValGluIleGly	155
DB	166	ACAGCTGCAGACAGATGGGCTATGACAGCCAGCCGCTTCAGAGCAGTGGAGATCCGT	225
QY	156	ProAspGlnAspLeuAspValValGluIleThrGluAsnSerGlnGluLeuArgMetArg	175
DB	226	CCAGATCAGAACTCCCTCTGCTCAAGTCACAGGAAACAGCAGGAACTTCAGGTGCAG	285
QY	176	AnSerSerGlyProCysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGly	195
DB	286	AATGGAGCAGATCTGCCTCTCAGGCTCCCTGGTTCTTCCGCTGCCTTGACTGTGA	345
QY	196	LysSerLeuLysThrProArgValValGlyGlyGluGluAlaSerValAspSerTrpPro	215
DB	346	AAGAGCTGAAGACTCTCTGCTGTGGTGTGGTGGGGTGGAGGCCCTCTGGATTCTTGGCCG	405
QY	216	TrpGlnValSerIleGlnTyrAspLysGlnHisValCysGlyGlySerIleLeuAspPro	235

```

Db      406 TGGCAGGTCCAGATCCAGTACCAAGCAGCATGCTGTGTGGTGGAGCATCTGGATCCC 465
Qy      236 HisTrpValLeuThrAlaAlaHisCysPheArgLysHisThrAspValPheAsnTrpLys 255
Db      466 CACTGGATCCTCACAGCAGCCACTGCTTCAGGAGTATCTTGATGTCAAGCTGGAAG 525
Qy      256 ValArgAlaGlySerAspLysLeuGlySerPheProSerLeuAlaValAlaLysIle 275
Db      526 GTACGGGAGGCTCAACATACCTGGTAACCTCTCCATCTGCTGTGGCCAAGATCTTC 595
Qy      276 IleIleGluPheAsnProMetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGln 295
Db      586 ATCGTGAACCATCTCTGTATCCCAAGAGAGACATTCCTCTGTTAAGTGCAG 645
Qy      296 PheProLeuThrPheSerGlyThrValArgProIleCysLeuProPhePheAspGluGlu 315
Db      646 ATGCCACTCACATCTCAGGCTCAGTCAGGCCCATCTGCTGCTCTCTGATGAGGTG 705
Qy      316 LeuThrProAlaThrProLeuTrpIleIleGlyTrp-GlyPheThrLysGlnAsn-GlyG 335
Db      706 CTGTGTCACGCCACACCATGCTGGTTCATTGGATGGGGCTTTACAGAAAGAAACCGGAG 765
Qy      335 LysMetSerAspIleLeuLeuGln-AlaSerVal-GlnValIleAspSerThrArgCy 354
Db      766 GAAGATGCTCTGACATCTACTCAAGGCATCAGTCCAGAGTCATCGAACGCACACGGTG 825
Qy      354 s-AsnAlaAspAspAlaTyr---GlnGlyGluValThrGluLysMetMetCysAlaGlyI 373
Db      826 GCATGTCAGAGGATGCTACCCAAAGGGGAGGACCGCTCGAATGTGGGGGACAGG 885
Qy      373 lePro---GluGlyGlyValAspThr---CysGlnGlyAspSer---GlyGlyProLeuM 390
Db      886 TACCCACACAGGTGGGCGAAGGACACCTGCCAGGGGGGCCAGTGGCGGGGCCCTTT 945
Qy      390 et 390
Db      946 TG 947

CV570449 602 bp mRNA linear EST 22-OCT-2004
oe01c10.y1 Human keratoconus cornea, unamplified, (od/oe) Homo
CV570449
CV570449.1 GI:54470982
EST.
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 602)
Rabinowitz, Y., Dong, L. and Wistow, G.
Expressed sequence tag analysis of human keratoconus cornea
JOURNAL
Unpublished (2004)
COMMENT
Contact: Wistow G
Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 496 0078
Email: graeme@helix.nih.gov
Plate: 01 row: c column: 10
Seq primer: M13RP1 reverse primer (ABI).
Location/Qualifiers
1. 602
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="oe01c10"
/tissue_type="Cornea"
/dev_stage="Adult"

FEATURES
source
RESULT 11
CV570449
LOCUS
DEFINITION
ACCESSION
CV570449
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
Rabinowitz, Y., Dong, L. and Wistow, G.
TITLE
Expressed sequence tag analysis of human keratoconus cornea
JOURNAL
Unpublished (2004)
COMMENT
Contact: Wistow G
Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 496 0078
Email: graeme@helix.nih.gov
Plate: 01 row: c column: 10
Seq primer: M13RP1 reverse primer (ABI).
Location/Qualifiers
1. 602
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="oe01c10"
/tissue_type="Cornea"
/dev_stage="Adult"

RESULTS
BP253625 Sugano cDNA library, kidney epithelial cell Homo sapiens
BP253625
LOCUS
DEFINITION
CDNA clone HRC04440, mRNA sequence.

```

```

/lab_host="EMDH10B"
/clone_lib="Human keratoconus cornea, unamplified,
(od/oe)"
/note="Organ: Eye; Vector: pCMVSPORT6; Approximately 40ug
total RNA was extracted from 7 adult human keratoconus
corneas. A directionally cloned cDNA library in the
pSPORT1 vector (Invitrogen) was constructed at Bioserve
Biotechnology (Laurel MD) essentially following the
protocols of the SuperScript plasmid system full details
of which are contained in the manufacturer's instruction
manual (http://www.lifetech.com/). First strand synthesis
was carried out using a Not I primer-adaptor
[5'-pGACTAGTCTAGATCGGAGCGGCC(T)15-3']. cDNA was
cloned in Not I/Sal I sites. EST analysis was performed on
the unamplified library at the NIH Intramural Sequencing
Center (NISC)."
```

ORIGIN

```

Alignment Scores:
Pred. No.: 1.22e-98 Length: 602
Score: 1037.00 Matches: 193
Percent Similarity: 98.97% Conservative: 0
Best Local Similarity: 98.97% Mismatches: 0
Query Match: 44.28% Indels: 2
DB: 7 Gaps: 0
```

US-10-803-530-2 (1-435) x CV570449 (1-602)

```

Qy      243 HisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAspLys 262
Db      1 CACTGTTCTCAGGAAACATACCGATGTGTTCACTGGAGGTGGGCGAGGCTCAGACAAA 60
Qy      263 LeuGlySerPheProSerLeuAlaValAlaLysIleIleIleIleIleIleIleIleIle 282
Db      61 CTGGGCGAGCTTCCCATCTCCCTGCTGTGGCCAAAGATCATCATTTGAATTCACCCCATG 120
Qy      283 TyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSerGly 302
Db      121 TACCCCAAGACATGACATCGCCCTCATGAAGCTGCAGTTCCCACTCACTTTCTCAGGC 180
Qy      303 ThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrProLeu 322
Db      181 ACAGTCAGGCCCATCTGTCTGCCCTCTTTGATGAGGAGCTCACTCCAGACCCACCCATC 240
Qy      323 TrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeuLeu 342
Db      241 TGGATCATTTGGATGGGGCTTTACGAAGCAGAAATGGAGGAAAGATGTTCTGACATCTG 300
Qy      343 GlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGlnGly 362
Db      301 CAGGCGTCAGTCCAGGTCAATTGACAGCACACGGTGCAATGCAGACGATCGGTACAGGG 360
Qy      363 GluValThrGluLysMetMetCysAla-GlyIleProGlu-GlyGlyValAspThrCysG 382
Db      361 GAAGTCACGAGAGAGATGATGTGCAGGGGCATCCCGAAGGGGGGTGTGGACACCTGCC 420
Qy      382 InGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIleV 402
Db      421 AGGGTGACAGTGGTGGGGCTGATGTACCAATCTGACCAAGTGGCATGTGGTGGGCATCG 480
Qy      402 alSerTrpGlyTyrGlyCysGlyProSerThrProGlyValTyrThrLysValSerA 422
Db      481 TTATGTGGGCTATGGCTGGGGGGCCGAGCACCAGGAGTATACACCAAGTCTCAG 540
Qy      422 laTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db      541 CCTATCTCACTGGATCTACATGTCTGGAAGGCTGAGCTG 581
```

```

RESULT 12
BP253625
LOCUS
DEFINITION
```

```
ACCESSION BP253625
VERSION BP253625.1 GI:52135906
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 592)
AUTHORS Suzuki, Y., Yamashita, R., Shirota, M., Sakakibara, Y., Chiba, J.,
Mizushima-Sugano, J., Nakai, K. and Sugano, S.
TITLE Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES
source
1..592
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HRC04440"
/tissue_type="kidney"
/cell_type="epithelial cell"
/clone_lib="Sugano cDNA library, kidney epithelial cell"
ORIGIN
Alignment Scores:
Pred. No.: 3 49e-97 Length: 592
Score: 1023.00 Matches: 193
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 43.68% Indels: 0
DB: 3 Gaps: 0
US-10-803-530-2 (1-435) x BP253625 (1-592)
QY 38 LeuLeuSerLeuAlaSerIlelleValValValLeuIleLysValIleLeuAspLys 57
Db 3 CTACTGAGCTGGCGAGTATCATCTGTTGTTCTCATCAGGTGATCTGGATAA 62
QY 58 TyrTyrPheLeuCysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGly 77
Db 63 TACTACTTCTCTGGGGCAGCTCTCCACTTCATCCCGAAGCAGCTGTGTGACGA 122
QY 78 GluLeuAspCysProLeuGlyGluAspGluGluHisCysValLysSerPheProGluGly 97
Db 123 GAGCTGGACTGTCCCTTGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTCCCGAAGG 182
QY 98 ProAlaValAlaValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAla 117
Db 183 CTGTGAGTGGCAGTCCGCTCTCCAGGACCGATCCACTGCAGGTGCTGGACTCGGC 242
QY 118 ThrGlyAsnTrpPheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAla 137
Db 243 ACAGGAACTGGTCTCTCGCTGTTTCGACAACTTCACAGAAGCTCTCGCTGAGACGCC 302
QY 138 CysArgGlnMetGlyTyrSerSerLysProThrPheArgAlaValGluLeuGlyProAsp 157
Db 303 TGTAGGCAGATGGCTACAGCAGCAAAACCCACTTTCAGAGCTGTGGAGATTGCCCGAC 362
QY 158 GlnAspLeuAspValValGluIleThrGluAsnSerGlnLeuLeuArgMetArgAnSer 177
Db 363 CAGGATCTGGATGTGTGTAATACAGAAACAGCGAGGAGCTTCGCATGCGGAATCA 422
QY 178 SerGlyProCysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSer 197
Db 423 AGTGGGCCCCCTCTCTCAGGCTCCCTCGGTCTCCCTGCACCTGTCTTGCTGTGGAAAGCC 482
QY 198 LeuLysThrProArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGln 217
Db 483 CTGAAGACCCCCCTGTGTGGTGGGGAGAGAGCCCTCTGTGGATTCTTGGCCTTGGCAG 542
QY 218 ValSerIleGlnTyrAspLysGlnHisValCysGlyGly 230
Db 543 GTCAGCATCCAGTACGACAAACAGCAGCTCTGTGGAGG 581
RESULT 13
CV028368 591 bp mRNA linear EST 20-AUG-2004
6865 Full Length cDNA from the Mammalian Gene Collection Homo
sapiens cDNA 5' similar to BC011703, mRNA sequence.
CV028368
ACCESSION CV028368.1 GI:51486447
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 591)
AUTHORS Rual, J.F., Hirozane-Kishikawa, T., Hao, T., Bertin, N., Li, S.,
Dricot, A., Li, N., Rosenberg, J., Lamesch, P., Vidalain, P.O.,
Clingingsmith, T.R., Hartley, J.L., Eposito, D., Cheo, D., Moore, T.,
Simmmons, B., Sequerra, R., Bosak, S., Doucette-Stamm, L., Le Peuch, C.,
Vandenhaute, J., Cusick, M.E., Albaladejo, J.S., Hill, D.E. and Vidal, M.
TITLE Human ORFome Version 1.1: a platform for Reverse Proteomics
JOURNAL Genome Res. (2004) In press
COMMENT Contact: Vidal M
Marc Vidal Laboratory
Dana Farber Cancer Institute
1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA
Tel: 617 632 5180
Fax: 617 632 5739
Email: Marc_Vidal@dfci.harvard.edu
ORF Sequence Tag (OST) of Gateway Entry construct. Each cloned ORF
results from a PCR reaction using an MGC full-length cDNA as
template DNA and ORF specific primers
PCR Primers
FORWARD: ATGTTACAGGATCTCTGACAGTGATCAA
BACKWARD: TACAGCTCAGCCTTCACAGCATTT
Insert Length: 591 Std Error: 7.00
Plate: 11036 row: 09 column: G
Seq primer: ACTGGCGTGTGTTTACAACTGTGACTGGGAAAC
High quality sequence start: 100
High quality sequence stop: 590
POLYA=No.
FEATURES
Location/Qualifiers
1..591
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="mixed"
/clone_lib="Full Length cDNA from the Mammalian Gene
Collection"
/note="Vector: mixed; The ORFs were PCR amplified from the
MGC (Mammalian Gene Collection) as of April 2004 and
cloned by recombinational Gateway cloning into pDONR223
Donor vector. Reference : MGC (Mammalian Gene Collection)
Program Team, Generation and Initial Analysis of more than
15,000 Full-Length Human and Mouse cDNA Sequences. PNAS,
2002, 99(26), 16899-16903"
ORIGIN
Alignment Scores:
Pred. No.: 2 49e-96 Length: 591
Score: 1015.00 Matches: 194
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 43.34% Indels: 0
DB: 7 Gaps: 0
```

US-10-803-530-2 (1-435) x CV028368 (1-591)

```
QY 2 AspProAspSerAspGlnProLeuSenSerLeuAppValIysProLeuArgIysProArg 21
Db 10 GATCTCTGACAGTATCAACCTCTGACAGCCTCGATGTCAAAACCCCTCGCAACCCCGT 69
QY 22 IleProMetGluThrPheArgIysValGlyIlePheIleIleAlaLeuLeuSerLeu 41
Db 70 ATCCCATGAGACCTTCAGAAAGTGGGATCCCATCATCATAGCACTACTGAGCCCTG 129
QY 42 AlaSerIleIleIleValValLeuIleIleIleValIleLeuAspIysIysIysPheLeu 61
Db 130 GCGAGTATCATATTGTGTCTCTCATCAAGGTGATCTCGATAAATACTACTCTCTC 189
QY 62 CysGlyGlnProLeuHisPheIleProArgIysGlnLeuCysAspGlyGluLeuAspCys 81
Db 190 TCGGGGACGCTCTCCACTTCATCCGAGGAGCAGCTGTGTGACGAGAGCTGGACTGT 249
QY 82 ProLeuGlyGluAspGluHisCysValIysSerPheProGluGlyProAlaValAla 101
Db 250 CCCTTGGGGAGAGCAGAGGACCTGTGTCAAGAGCTTCCCGAAGGGCCCTGCAAGTGCA 309
QY 102 ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp 121
Db 310 GTCCGCTCTCCAAAGACCGATCCACACTGAGAGTGTGGACTCGGCCACAGGGAAGCTGG 369
QY 122 PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
Db 370 TTCTCTGCTGTGTTCGACAACTTCACAGAAGCTCTCGCTGAGACAGCCTGTAGGCAGATG 429
QY 142 GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
Db 430 GGCTACAGCAGCAAAACCCACTTTCAGAGCTGTGGAGATTGGCCACAGACCAAGATCTGGAT 489
QY 162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
Db 490 GTTGTGTAATCACAGAAACAGCCAGAGCTTCGATGCGGAATTCAGTGGGCCCTGTGT 549
QY 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGly 195
Db 550 CTCTCAGGCTCCCTGGTCTCGCTGCATGTCTTGCTGTGGG 591
```

```
RESULT 14
BP260542 573 bp mRNA linear EST 16-SEP-2004
LOCUS BP260542 Sugano cDNA library, small intestine Homo sapiens CDNA
DEFINITION clone HSI00780, mRNA sequence.
ACCESSION BP260542
VERSION BP260542.1 GI:52175773
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 573)
AUTHORS Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.
FEATURES
Location/Qualifiers
1..573
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HSI00780"
```

/tissue type="small intestine"
/clone_lib="Sugano cDNA library, small intestine"

```
ORIGIN
Alignment Scores: 1.02e-95 Length: 573
Pred. No.: 1009.00 Matches: 187
Score: 98.42% Conservative: 0
Percent Similarity: 98.42% Mismatches: 3
Best Local Similarity: 43.08% Indels: 0
Query Match: 3 Gaps: 0
DB: 3
US-10-803-530-2 (1-435) x BP260542 (1-573)
QY 172 LeuArgMetArgAsnSerSerGlyProCysLeuSerGlySerLeuValSerLeuHisCys 191
Db 2 CTTTGCATGCGGAACCTCAAGTGGCCCTGTCTCTCAGGCTCCCTGGTCTCCCTGCATGT 61
QY 192 LeuAlaCysGlyIysSerLeuLysThrProArgValValGlyGlyGluAlaSerVal 211
Db 62 CTTGCCTGTGGGAAGACCTGAAGACCCCGTGTGTGTGGTGGGAGAGGCTCTGTGT 121
QY 212 AspSerTrpProTrpGlnValSerIleGlnIysAspIysGlnHisValCysGlyGlySer 231
Db 122 GATTCTTGGGCTTGGCAGGTGAGCATCTAGTACGACAAACAGCACGCTCTGTGGAGGAGC 181
QY 232 IleLeuAspProHisTrpValLeuThrAlaAlaHisCysPheArgIysHisThrAspVal 251
Db 182 ATCTGAGACCCCTGCTGGTCTCAGGACGCCACTGTCTCAGGAACATACCGATGTG 241
QY 252 PheAsnTrpIysValArgAlaGlySerAspLysLeuGlySerPheProSerLeuAlaVal 271
Db 242 TTCAACTGGAAGTGGGGCAGGCTCAGACAAATCGGCAGCTTCCCATCCCTGGCTGTG 301
QY 272 AlaIysIleIleIleIleGluPheAsnProMetTyrProLysAspAsnAspIleAlaLeu 291
Db 302 GCCAAGATCATCATCATATTGAATTCACCCCATGTACCCCAAGACAAATGACATCGCCCTC 361
QY 292 MetLysLeuGlnPheProLeuThrPheSerGlyThrValArgProIleCysLeuProPhe 311
Db 362 ATGAGCTGCAGTTCCTCCACTCATCTTCTCAGGCACAGTCAGGCCCATCTGTCTGCCCTTC 421
QY 312 PheAspGluGluLeuThrProAlaThrProLeuTrpIleIleGlyTrpGlyPheThrLys 331
Db 422 TTTGATGAGAGGAGTCACTCCAGCCACCCCTCATCTGGATCATTGATGGGGCTTTACGAAG 481
QY 332 GlnAsnGlyGlyLysMetSerAspIleLeuLeuGlnAlaSerValGlnValIleAspSer 351
Db 482 CAGAAATGGAGGGAAGATGTCTGACATCTGCTGCACGCTGCAGCTCAGGTTCATTGACAGC 541
QY 352 ThrArgCysAsnAlaAspAlaTyrGln 361
Db 542 ACACGGTGCATCTGCAGACCGATGGTACCAG 571
```

```
RESULT 15
BP253686 583 bp mRNA linear EST 15-SEP-2004
LOCUS BP253686 Sugano cDNA library, kidney epithelial cell Homo sapiens
DEFINITION CDNA clone HRC04627, mRNA sequence.
ACCESSION BP253686
VERSION BP253686.1 GI:52135967
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 583)
AUTHORS Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
```

PUBMED
15342556

COMMENT
Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp.

FEATURES

source
1..583
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HRC04627"
/tissue_type="kidney"
/cell_type="epithelial cell"
/clone_lib="Sugano cDNA library, kidney epithelial cell"

ORIGIN

Alignment Scores:
Pred. No.: 1.93e-95 Length: 583
Score: 1006.50 Matches: 194
Percent Similarity: 97.49% Conservative: 0
Best Local Similarity: 97.49% Mismatches: 0
Query Match: 42.98% Indels: 5
DB: 3 Gaps: 1

US-10-803-530-2 (1-435) x BP253686 (1-583)

QY	19	LysProArgIleProMetGluThrPheArgLysValGlyIleProIleIleAlaLeu	38
DB	2	AAACCCCGATCCCATGAGACCTTCAGAAAGTGGGATCCCATCATAGCACTA	61
QY	39	LeuSerLeuAlaSerIleIleValValLeuIleLysValIleLeuAspLysTyr	58
DB	62	CTGAGCCTGGCGAGTATCATTTGTTGTCTCATCAAGGTGATCTGGATAATAC	121
QY	59	TyrPheLeuCysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGlu	78
DB	122	TACTTCTCTGGCGGAGCCTCTCCACTTCATCCCGAGGAAGCAGCTGTGTGACGGAGAG	181
QY	79	LeuAspCysProLeuGluGluAspGluGluHisCysValLysSerPheProGluGlyPro	98
DB	182	CTGGACTGTCCCTTGGGGGAGGACGAGGAGCACTGTCAAGAGCTTCCCGAAGGGCCT	241
QY	99	AlaValAlaValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThr	118
DB	242	GCAGTGGCAGTCCGCTCTCCAGGACCGATCCACACTGCAGGTGCTGGACTCGGCCACA	301
QY	119	GlyAsnTrpPheSerAlaCysPheAspAsnPhetrGluAlaLeuAlaGluThrAlaCys	138
DB	302	GGGAACCTGGTTCTCTGCTGTTTCGACAACTTCACAGAAGCTCTCGCTGAGACAGCCTGT	361
QY	139	ArgGlnMetClyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGln	158
DB	362	AGGCAGATGGGCTACAGC-----AGAGCTGTGAGATGGCCAGACCAG	406
QY	159	AspLeuAspValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSer	178
DB	407	GATCTGGATGTTGTTGNAATCACAGAAAACAGCCAGGAGCTTCGCATGCGGAACCTCAAGT	466
QY	179	GlyProCysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeu	198
DB	467	GGGCCCTGTCTCTCAGGCTCCCTGGTCTCCCTGCACGTCTTGCCTGTGGGAAGAGCCTG	526
QY	199	LysThrProArgValValGlyGluGluAlaSerValAspSerTrpProTrpGln	217
DB	527	AAGACCCCGCTGTGTGGTGGTGGAGGAGGCGCTCTGTGATTCTTTGGCCTTGGCAG	583

Search completed: December 6, 2005, 23:06:42
Job time : 4200 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 6, 2005, 23:06:51 ; Search time 238 Seconds
(without alignments)
568.827 Million cell updates/sec

Title: US-10-803-530-2
Perfect score: 2342
Sequence: 1 MDPDSQPLNSLDVPLRKP.....VYTKVSAYLNWYVWKAEL 435

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3289935 seqs, 155610033 residues

Total number of hits satisfying chosen parameters: 6579870

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2.1/USPTO.spool.p/US10803530/runat 05122005 083218 2272/app query.fasta_1.583
-DB=Published Applications NA.New -OFMT=fastap -SUFFIX=rnbpn -MINMATCH=0.1
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -MMAP=US10803530 @CGN 1 1 184 @runat 05122005 083218 2272
-NCPU=6 -ICPU=3 -NO USR= -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA.New:

- 1: /cgn2.6/ptodata/1/pubpna/US09 NEW PUB.seq.*
- 2: /cgn2.6/ptodata/1/pubpna/US06 NEW PUB.seq.*
- 3: /cgn2.6/ptodata/1/pubpna/US07 NEW PUB.seq.*
- 4: /cgn2.6/ptodata/1/pubpna/US08 NEW PUB.seq.*
- 5: /cgn2.6/ptodata/1/pubpna/PCT NEW PUB.seq.*
- 6: /cgn2.6/ptodata/1/pubpna/US10 NEW PUB.seq.*
- 7: /cgn2.6/ptodata/1/pubpna/US11 NEW PUB.seq.*
- 8: /cgn2.6/ptodata/1/pubpna/US11 NEW PUB.seq2.*
- 9: /cgn2.6/ptodata/1/pubpna/US11 NEW PUB.seq3.*
- 10: /cgn2.6/ptodata/1/pubpna/US60 NEW PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2297.5	98.1	2063	US-11-102-240-111	Sequence 111, App
2	584.5	25.0	1783	US-11-112-908-14	Sequence 14, Appl
3	580	24.8	1615	US-11-182-752-1	Sequence 1, Appli
4	488.5	20.9	4804	US-11-067-811-3	Sequence 3, Appli
5	475	20.3	2103	US-11-102-240-105	Sequence 105, App
6	473.5	20.2	918	US-11-137-465-23	Sequence 23, Appl
7	467.5	20.0	1129	US-10-131-826A-221	Sequence 221, App
8	449	19.2	1802	US-10-821-234-191	Sequence 191, App

9	442.5	18.9	780	7	US-11-056-621-5	Sequence 5, Appli
10	441.5	18.9	2376	6	US-11-056-621-3	Sequence 3, Appli
11	377.5	16.1	1204	6	US-10-131-826A-505	Sequence 505, App
12	369.5	15.8	2016	7	US-11-029-003-7	Sequence 7, Appli
13	356	15.2	1475	6	US-10-623-155-122	Sequence 122, App
14	356	15.2	2294	6	US-10-623-155-123	Sequence 123, App
15	355.5	15.2	2091	7	US-11-029-003-9	Sequence 9, Appli
16	346.5	14.8	999	6	US-10-131-826A-395	Sequence 395, App
17	319	13.6	1006	6	US-10-821-234-422	Sequence 422, App
18	313	13.4	711	6	US-10-401-386B-23	Sequence 23, Appl
19	307.5	13.1	2067	6	US-10-821-234-681	Sequence 681, App
20	306	13.1	1570	6	US-10-131-826A-455	Sequence 455, App
21	305.5	13.0	7681	7	US-11-094-484-2	Sequence 2, Appli
22	303	12.9	711	6	US-10-401-386B-5	Sequence 5, Appli
23	299.5	12.8	2187	7	US-11-094-484-1	Sequence 1, Appli
24	299.5	12.8	5181	7	US-11-094-484-3	Sequence 3, Appli
25	297	12.7	2118	6	US-10-821-234-680	Sequence 680, App
26	291	12.4	768	7	US-11-147-047-10	Sequence 10, Appl
27	289	12.3	1394	6	US-10-401-386B-3	Sequence 3, Appli
28	287.5	12.3	967	6	US-10-401-386B-1	Sequence 1, Appli
29	287.5	12.3	4775	6	US-10-401-386B-62	Sequence 62, Appl
30	285.5	12.2	585	6	US-10-401-386B-11	Sequence 11, Appl
31	284	12.1	1729	6	US-10-401-386B-4	Sequence 4, Appli
32	256.5	11.0	596	7	US-11-108-173-931	Sequence 931, App
33	239	10.2	150468	7	US-11-112-908-56	Sequence 56, Appl
34	239	10.2	193789	7	US-11-112-908-55	Sequence 55, Appl
35	231.5	9.9	906	7	US-11-147-047-11	Sequence 11, Appl
36	230	9.8	2501	6	US-10-821-234-182	Sequence 182, App
37	226	9.6	2846	7	US-11-103-240-37	Sequence 37, Appl
38	217	9.3	597	6	US-10-401-386B-28	Sequence 28, Appl
39	213.5	9.1	21784	7	US-11-182-752-3	Sequence 3, Appli
40	208	8.9	726	7	US-11-137-465-5	Sequence 5, Appli
41	208	8.9	732	7	US-11-137-465-6	Sequence 6, Appli
42	208	8.9	1123	6	US-10-131-826A-247	Sequence 247, App
43	204.5	8.7	1654	6	US-10-401-386B-7	Sequence 7, Appli
44	184.5	7.9	601	7	US-11-182-752-14	Sequence 14, Appl
45	173.5	7.4	1496	6	US-10-750-185-47031	Sequence 47031, A

ALIGNMENTS

RESULT 1

US-11-102-240-111
; Sequence 111, Application US/11102240
; Publication No. US20050260647A1
; GENERAL INFORMATION:
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: ANTIBODIES TO POLYPEPTIDES ENCODED BY A NUCLEIC ACID UNDEREXPRES
; FILE REFERENCE: P3230RIC106C
; CURRENT APPLICATION NUMBER: US/11/102,240
; CURRENT FILING DATE: 2003-04-08
; PRIOR APPLICATION NUMBER: 10/063662
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 10/006867
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 60/170262
; PRIOR FILING DATE: 199-12-09
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 111
; LENGTH: 2063
; TYPE: DNA
; ORGANISM: Homo Sapien
US-11-102-240-111

Alignment Scores: 1.13e-243 Length: 2063
Pred. No.: 1.13e-243

```
Score: 2297.50 Matches: 429
Percent Similarity: 98.85% Conservative: 0
Best Local Similarity: 98.85% Mismatches: 0
Query Match: 98.10% Indels: 5
DB: 7 Gaps: 1

US-10-803-530-2 (1-435) x US-11-102-240-111 (1-2063)

QY 2 AspProAspSerGlnProLeuAsnSerLeuAspValLysProLeuArgLysProArg 21
Db 219 GATCCTGACAGTGATCAACCTCTGAACAGCCTCGAATGTCAAACCCCTGGCGAAACCCCGT 278
QY 22 IleProMetGluThrPheArgLysValGlyLeuProIleIleIleAlaLeuLeuSerLeu 41
Db 279 ATCCCCATGAGACCTTCAGAAAGTGGGATCCCCCATCATAGCACTACTGAGCCTG 338
QY 42 AlaSerIleIleIleValValLeuIleValLysValLysValLysValLysValLysVal 61
Db 339 GCGAGTATCATCATTTGGTGTGCTTCATCAAGGTGATTTCTGGATAAATACTACTTCCTC 398
QY 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys 81
Db 399 TCGCGCAGCCTCTCCATTCATCCGAGAGAGCAGCTGTGTGACGGAGAGCTGGACTGT 458
QY 82 ProLeuGlyGluAspGluGluHisCysValLysSerPheProGluGlyProAlaValAla 101
Db 459 CCCCTGGGGGAGGAGGAGGAGCACTGTGTCAAGAGCTTCCCCGAAGGGCTTCGAGTGCA 518
QY 102 ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp 121
Db 519 GTCCGCTCTCCAAGGACCGATCCACATGCGAGTGTGGACTCGGCCACAGGGAACCTGG 578
QY 122 PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
Db 579 TTCTCTGCTGTGTTCGACAACTTCACAGAGCTCTCGTGAGACGCTGTAGCAGATG 638
QY 142 GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
Db 639 GCGTACAGC-----AGAGCTGTGGAGATTGGCCCCCAGACCAAGATCTGGAT 683
QY 162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
Db 684 GTTGTGTAATCACAGAAACAGCCAGGAGCTTCGCATGCGGAATCAAGTGGGCCCTGT 743
QY 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro 201
Db 744 CTCCTAGGCTCCCTGGTCTCCCTGCACTGTCTGCTGTGGAAAGAGCCTGAAGACCCCC 803
QY 202 ArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
Db 804 CGTGTGGTGGTGGGAGGAGCCTCTGTGGATTCTTGGCCTTGGCAGGTCAGATCCAG 863
QY 222 TyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAla 241
Db 864 TACACAAACAGCAGCTGTGTGGAGGAGCATCTGGACCCCCACCTGGGCTCTCACGGCA 923
QY 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
Db 924 GCGCAGCTGCTTCAGAAACATACCCGATGTGTCAACTGGAAGGTGGCGGAGGCTCAGAC 983
QY 262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleIleGluPheAsnPro 281
Db 984 AAATGGGAGCTTCCATCCCTGGCTGTGGCCAGAGATCATCATTTGAATTAACACCCC 1043
QY 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db 1044 ATGTACCCCAAGACAATGATGCGCCCTCATGAAGCTGCAGTTCACCTACTCTCTCA 1103
QY 302 GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro 321
Db 1104 GGCAGCTACGGCCCATCTCTCTGCCCTCTTTGTATGAGGAGCTACTCCAGCCACCCCA 1163
QY 322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu 341
```

```
Db 1164 CTCCTGGATCATTTGGATGGGCTTTACGAAGCAGAATGGAGGAAGATGCTCTGACATCTG 1223
QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGln 361
Db 1224 CTGCAGGCGTCAGTCCAGGTCAITGACAGCACCGGTGCAATGCAGACGATGCGTACCCAG 1283
QY 362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys 381
Db 1284 GGGGAAGTCAACCGAGAAGATGATGTGTGAGGATCCCGGAAGGGGTGTGGACACCTGC 1343
QY 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle 401
Db 1344 CAGGTGACAGTGTGTGGCCCCCTGATGTACCAATCTGACCAGTGCATGTGTGGGCGATC 1403
QY 402 ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSer 421
Db 1404 GTTAGCTGGGCTATGGCTGCGGGGGCCCGAGCACCCAGGAGTATACCAAGGTCTCA 1463
QY 422 AlaTyrIleAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db 1464 GCCTATCTCAACTGGATCTACAATGTCTGGAAGGCTGAGCTG 1505

RESULT 2
US-11-112-908-14
; Sequence 14, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Lisa M.
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT APPLICATION NUMBER: US/11/112,908
; PRIOR FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/564,758
; PRIOR FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/575,978
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/631,702
; PRIOR FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: US 60/633,826
; PRIOR FILING DATE: 2004-12-07
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 14
; LENGTH: 1783
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-112-908-14

Alignment Scores:
Pred. No.: 3 81e-54 Length: 1783
Score: 584.50 Matches: 140
Percent Similarity: 45.0% Conservative: 60
Best Local Similarity: 31.53% Mismatches: 161
Query Match: 24.96% Indels: 83
DB: 7 Gaps: 11

US-10-803-530-2 (1-435) x US-11-112-908-14 (1-1783)
QY 19 LysProArgIleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeu 38
Db 288 AGACCAAGGTGGCAGCTCTCACT-----GGGGGACCCCTGCTACTTCTGACAGCC 338
QY 39 LeuSerLeuAlaSerIleIleIleValValLeuIleLysValLysValLysValLysVal 58
Db 339 ATCGGGGCGGATCTCTGGCCATTTGTGGCTGTCTCTCAGG----- 380
QY 59 TyrPheLeuCysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGlu 78
Db 380 ----- 380
QY 79 LeuAspCysProLeuGlyGluAspGluHisCysValLysSerPheProGluGlyPro 98
```



```
Db 381 -----AGTACCAGGAG-----CCG 395
QY 99 AlaValAlaValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThr 118
Db 396 CTGTACCCAGTCAGTCAGTCCTGCGGACGCTCGGCTCATGCTCTTTGACACGCGAA 455
QY 119 GlyAsnTrpPheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCys 138
Db 456 GGGACGTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 515
QY 139 ArgGlnMetGlyTyrSerSerLysProThrPheArgAlaValAluIleGlyProAspGln 158
Db 516 GAGGAGTGGCTTC-----CTCAGGGCACTG-----ACCCACTCC 551
QY 159 AspLeuAspValValGluIleThrGluAsn-----168
Db 552 GAGCTGGACGTGCAACCGCGCGCGCAATGGCAGCTCGGCTTCTTCTGTGGACGAG 611
QY 169 -----SerGlnGluLeuArgMetArgAsnSerSerGlyProCysLeuSer 183
Db 612 GGGAGGCTCCCCACACCCAGAGGCTGCTGGAGGTCATCTCGGTGTGATTGCCCCAGA 671
QY 184 GlySerLeuValSerLeuHisCysLeuAlaCysGly---LysSerLeuLysThrProArg 202
Db 672 GCGCGTTCTTGGCGCCATCTGCCAGACTGTGGCGCGAGGAGCTGCCGTGGACCGC 731
QY 203 ValValGlyGlyGluAlaSerValAspSerTrpProTrpGlnValSerIleGlnTyr 222
Db 732 ATCTGGTGGAGGCGGACACACAGCTTGGCGCGGTGGCGTGCAGTCACTGCTGCTAT 791
QY 223 AspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAlaAla 242
Db 792 GATGAGACACCTCTGTGGGGGATCCCTGCTCTCGGGGAGTGGGTGTGACAGCGGC 851
QY 243 HisCysPheArgLysHisThrAspValPheAsn---TrpLysValArgAlaGlySerAsp 261
Db 852 CACTCTTCCGAGCGGACCGGTCCTGTCCGATGGAGTGTTCGCGTGGCGTG 911
QY 262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsnPro 281
Db 912 GCCAGGCGCTCTCCCGACCGTCTGCAGCTGGGGTGCAGGCTGTGCTACCCAGCGGC 971
QY 282 MetTyrPro-----LysAspAsnAspIleAlaLeuMetLys 293
Db 972 TATCTTCCCTTTTCGGACCCCAACAGCAGGAGAACAGCAAGATATTGCCCTGTGCCAC 1031
QY 294 LeuGlnPheProLeuThrPheSerGlyThrValArgProIleCysLeuProPheAsp 313
Db 1032 CTCTCCAGTCCCTGCCCCCTCACAGAAATACATCCAGCCTGTGTGCTCCAGTCCCGC 1091
QY 314 GluGluLeuThrProAlaThrProLeuTrpIleIleGlyTrpGlyPheThrLysGlnAsn 333
Db 1092 CAGGCGCTGTGGATGGCAAGATCTGACCGTACCGGGTGGGGCAACAG---CAGTAC 1148
QY 334 GlyGlyLysMetSerAspIleLeuLeuGlnAlaSerValGlnValIleAspSerThrArg 353
Db 1149 TATGCCAACAGCGCGGGTACTCCAGAGGCTCGAGTCCCATATCAGCAATGATGTC 1208
QY 354 CysAsnAlaAspAspAlaTyrGlnGlyGluValThrGluLysMetMetCysAlaGlyIle 373
Db 1209 TCAATATGGCGCTGACTTCTATGGAAACCAAGATCAAGCCCAAGATGTCTGTGCTGCTAC 1268
QY 374 ProGluGlyGlyValAspThrCysGlnGlyAspSerGlyGlyProLeuMetTyrGln--- 392
Db 1269 CCGAGGGTGGATTATGCTCCAGCGCGACAGCGGGTGGTCCCTTTGTGTGGAGAC 1328
QY 393 -----SerAspGlnTrpHisValValGlyIleValSerTrpGlyTyrGlyCys 408
Db 1329 AGCATCTCTCGGACGCCACGTTGGCGCTGTGTGGCATGTGAGTTGGGGCACTGGCTGT 1388
QY 409 GlyGlyProSerThrProGlyValTyrThrLysValSerAlaTyrLeuAsnTrpIleTyr 428
```

```
Db 1389 GCCCTGGCCAGAGCCAGGCGTCTACACCAAGTCAGTGACTTCCGGGAGTGATCTTC 1448
QY 429 AsnValTrpLys 432
Db 1449 CAGGCCATAAAG 1460
RESULT 3
US-11-182-752-1
; Sequence 1, Application US/11182752
; Publication NO. US20050250154A1
; GENERAL INFORMATION:
; APPLICANT: Gan, Weiniu et al.
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL001194DIVII
; CURRENT APPLICATION NUMBER: US/11/182,752
; CURRENT FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1615
; TYPE: DNA
; ORGANISM: HUMAN
US-11-182-752-1
Alignment Scores:
Pred. No.: 1,046-53 Length: 1615
Score: 580.00 Matches: 132
Percent Similarity: 44.06% Conservative: 57
Best Local Similarity: 30.77% Mismatches: 146
Query Match: 24.77% Indels: 94
DB: 7 Gaps: 9
US-10-803-530-2 (1-435) x US-11-182-752-1 (1-1615)
QY 19 LysProArgIleProMetClnThrPheArgLysValGlyIleProIleIleIleAlaLeu 38
Db 220 AGACCCAAAGTGGCAGCTCTCACT-----GCGGGGAGCCCTGCTACTTCTGACAGCC 270
QY 39 LeuSerLeuAlaSerIleIleIleValValLeuIleLysValIleLeuAspLysTyr 58
Db 271 ATCGGGGCGGCATCTCGGCCATTTGGCTGTCTCTCTCAGG-----312
QY 59 TyrPheLeuCysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGlu 78
Db 312 -----312
QY 79 LeuAspCysProLeuGlyGlnAspGluHisCysValLysSerPheProGluGlyPro 98
Db 313 -----AGTACCAGGAG-----CCG 327
QY 99 AlaValAlaValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThr 118
Db 328 CTGTACCCAGTCAGTCAGTCCTCGGACGCTCGGCTCATGCTCTTTCACACGCGAA 387
QY 119 GlyAsnTrpPheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCys 138
Db 388 GGGACGTGGCGGCTGCTGCTGCTCCTCGGCTCCCAAGCCAGGAGTACCGGACTCAGCTGC 447
QY 139 ArgGlnMetGlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGln 158
Db 448 GAGGAGTGGGCTCTCTCAGT-----468
QY 159 AspLeuAspValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSer 178
Db 468 -----468
QY 179 GlyProCysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGly---LysSer 197
Db 469 ---GATTGCCAGAGCGCGTTTCTGGCGGCCCATCTGCAAGACTGTGCGCCGAGGAAG 525
QY 198 LeuLysThrProArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGln 217
```

```
Db 526 CTGCGCGTGACCGCATCGTGGAGGCGGGACACACAGCTTGGGCGCGTGGCGCA 585
Qy 218 ValSerIleGlnTyrAspLysGlnHisValCysGlySerIleLeuAspProHisTrp 237
Db 586 GTCAGCTTCGCTATGATGAGACACACCTCTGTGGGGATCCCTGCTCCGGGAGCTGG 645
Qy 238 ValLeuThrAlaAlaHisCysPheArgLysHisThrAspValPheAsn---TrpLysVal 256
Db 646 GTGCTGACAGCGCCCACTGCTTCCGGAGCGGACCGGGTCTCTCCGATGGCGAGTG 705
Qy 257 ArgAlaGlySerAspLysLeuGlySerPheProSerLeuAlaValAlaLysIleIle 276
Db 706 TTTGCGGCTGCGTGGCCAGGCGCTCTCCCAAGTCTGCAGCTGGGGTGCAGGCTGTG 765
Qy 277 IleGluPheAsnProMetTyrPro-----LysAspAsnAsp 288
Db 766 GTCTACACCGGGGCTATCTTCCCTTCCGGACCCCAACAGCAGGAGAGACAGCAACGAT 825
Qy 289 IleAlaLeuMetLysLeuGlnPheProLeuThrPheSerGlyThrValArgProIleCys 308
Db 826 ATTGCCCTGGTCCACCTCCAGTCCCTGCGCTCACAGATACATCCAGCCTGTGTGC 885
Qy 309 LeuProPhePheAspGluGluLeuThrProAlaThrProLeuTrpIleIleGlyTrpGly 328
Db 886 CTCCAGCTGCGCGGCGGCGCTGTGGATGGCAAGATCTGTACCGTGCAGCGGCTGGGCG 945
Qy 329 PheThrLysGlnAsnGlyLysMetSerAspIleLeuLeuGlnAlaSerValGlnVal 348
Db 946 AACACG---CAGTACTATGGCCACACAGCGCGGGTACTCCAGGAGCTCGAGTCCCNATA 1002
Qy 349 IleAspSerThrArgCysAsnAlaAspAlaTyrGlnGlyGluValThrGluLysMet 368
Db 1003 ATCACCATGATGCTCGAATGGCGCTGACTTCTATGGAACACAGATCAAGCCCAAGATG 1062
Qy 369 MetCysAlaGlyIleProGluGlyGlyValAspThrCysGlnGlyAspSerGlyGlyPro 388
Db 1063 TTCTGTGCTGGTACCCCGAGGGTGGCATTGATGCTCCAGGCGGACAGCGGGTGGTCCC 1122
Qy 389 LeuMetTyrGln-----SerAspGlnTrpHisValValIleValSer 403
Db 1123 TTTGTGTGTAGGACAGCATCTCTCGGACGCCAGCTTGGCGGCTGTGTGGCATTGTGAGT 1182
Qy 404 TrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSerAlaTyr 423
Db 1183 TGGGCACTGCTGTGCGCTGGCCAGCAGAGCCAGGCGGTCTACACCAAGTCAGTGACTTC 1242
Qy 424 LeuAsnTrpIleTyrAsnValTrpLys 432
Db 1243 CGGAGTGGATCTTCCAGGCCATAAG 1269
```

RESULT 4

```
US-11-067-811-3
; Sequence 3, Application US/11067811
; Publication No. US2005026068A1
; GENERAL INFORMATION:
; APPLICANT: Morgan, Bruce A
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR HAIR GROWTH
; FILE REFERENCE: 10287-083001
; CURRENT APPLICATION NUMBER: US/11/067,811
; PRIOR FILING DATE: 2005-02-28
; PRIOR APPLICATION NUMBER: US 60/548,272
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3
; LENGTH: 4804
; TYPE: DNA
; ORGANISM: Mus Musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3339)
```

US-11-067-811-3

```
Alignment Scores:
Pred. No.: 6.92e-43 Length: 4804
Score: 488.50 Matches: 130
Percent Similarity: 46.55% Conservative: 52
Best Local Similarity: 33.25% Mismatches: 148
Query Match: 20.86% Indels: 61
Db: 7 Gaps: 17
```

US-10-803-530-2 (1-435) x US-11-067-811-3 (1-4804)

```
Qy 69 IleProArgLysGlnLeuCysAspGlyGluLeuAspCysProLeuGlyGluAspGluGlu 88
Db 2206 GTGCGCGGTGACCTTTGGTGGCAGCGATGGTGCAGCTGCTCAGACAGTCTTGATGAATGG 2265
Qy 89 HisCysValLysSerPheProGluGlyProAlaValAlaValArgLeuSerLysAspArg 108
Db 2266 GGCTGTGTGACCCCTCTCTAAATAATGG-----AACTCTCTCC 2301
Qy 109 SerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrpPheSerAlaCysPheAspAsn 128
Db 2302 TCATTGTGACTGTTCAAAATCTGCAAGGAACAC-----CACGTGTGTGCTGACGGC 2355
Qy 129 PheThrGluAlaLeuAlaGluThrAlaCysArgGlnMetGlyTyrSerSerLysProThr 148
Db 2356 TGGCGGAGAGCTTGAGTCHAGCTGGCTGCAAGCAGATGGTTTAGGAGAACCTCTCTGTG 2415
Qy 149 PheArgAlaValGluIleGlyProAspGlnAspLeuAspValValGluIleThrGluAsn 168
Db 2416 ACCAAGCTGATC-----CCAGCAGCAG-----GAAGGC 2442
Qy 169 SerGlnGluLeuArgMet-----ArgAsnSerSerGly----- 179
Db 2443 CAGCAGTGGCTGAGGTGTACCCCACTCGGAGAAATCTCAATGGGAGCACCTTGCAGGAG 2502
Qy 180 -----ProCysLeuSerGlySerLeuValSerLeuHisCysLeuAla 193
Db 2503 CTGCTGGTATACAGGCACCTCTGCCCAAGCAGAGTGAATTTCCCTTCTGTGCTCCAAG 2562
Qy 194 -----CysGlyLys-----SerLeuLysThrProArgValValGlyGlyGluGlu 208
Db 2563 CAAGACTGTGGCGCGCGCCCTGCTGCCGAATGAACAGAGGATCCTTGGGGTTCGGACT 2622
Qy 209 AlaSerValAspSerTrpProTrpGlnValSerIleGlnTyrAspLysGln---HisVal 227
Db 2623 AGTCGCTCGGAGGTGGCGGTGGCAGTCTCTCTGCAAGTGAACCCAGTCGACATATC 2682
Qy 228 CysGlyGlySerIleLeuAspProHisTrpValLeuThrAlaAlaHisCysPheArgLys 247
Db 2683 TGTGGCTGTCTCTCATTTGCCAAGAGTGGTCTCTGACAGTTGCCATTGCTTTGAAGGG 2742
Qy 248 HisThrAspValPheAsnTrpLysValArgAlaGlySerAspLysLeu-----Gly 264
Db 2743 AGAGAAGACGCTGATGTTTGGAAAGTGTATTTTGGCATAAACAACACCTGCACCATCA 2802
Qy 265 SerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsnProMetTyrPro 284
Db 2803 GCGTTTCATGCAGACCGCGTTTGTGAAGACCATCTCTGCTA-----CATCCCCGTACAGT 2856
Qy 285 Lys-----AspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db 2857 CGAGCAGTGGTAGACTATGATATCAGCGTGTGGAGCTGAGCGATGATATCAATGAGACA 2916
Qy 302 GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro 321
Db 2917 AGCTACGTCAGACCTGTCTGCTACCCAGTCCGGAGGAGTATCTAGAACCAACGATACGTAC 2976
Qy 322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu 341
Db 2977 TGCTACATCACAGGCTGGGGC-----CACATGGGCAATAAAATGCCCTTTAAGCTG 3027
Qy 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaTyr--- 360
```

```

Db 3028 CAGGAGGAGAGGTCGCGCATTTATCCCTCTGAGCAGTGC-----CAGTCCTATTTT 3078
Qy 361 ---GlnGlyGluValThrGluLysMetCysAlaGlyIleProGluGlyGlyValAsp 379
Db 3079 GACATGAAGACCATCACCAATCGATGATCTGCTGGCTATGATCTGGCAGCGTGGAC 3138
Qy 380 ThrCysGlnGlyAspSerGlyProLeuMetTyrGln-----SerAspGlnTrpHis 397
Db 3139 TCCTGCATGGGAGACAGCGGTGGGCTCTGGTTGTGAACACCGGAGGACAGTGGACA 3198
Qy 398 ValValGlyIleValSerTrpGlyTyrGlyCysGlyGlyProSerThr---ProGlyVal 416
Db 3199 TTATTTGGTTAACTTCATGGGGCTCCGCTGCTTTTCCAAAGTTCTGGGACCTGGAGTG 3258
Qy 417 TyrThrLysValSerAlaTyrLeuAsnTrpIle 427
Db 3259 TACAGCAATGTCTTACTTTTGGGGCTGGATT 3291

RESULT 5
US-11-102-240-105
; Sequence 105, Application US/111102240
; Publication No. US20050260647A1
; GENERAL INFORMATION:
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: ANTIBODIES TO POLYPEPTIDES ENCODED BY A NUCLEIC ACID UNDEREXPRESS
; FILE REFERENCE: P3230R1C106C
; CURRENT APPLICATION NUMBER: US/11/102,240
; CURRENT FILING DATE: 2005-04-08
; PRIOR APPLICATION NUMBER: 10/063662
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 10/006867
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 60/170262
; PRIOR FILING DATE: 199-12-09
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 105
; LENGTH: 2103
; TYPE: DNA
; ORGANISM: Homo Sapien
US-11-102-240-105

Alignment Scores:
Pred. No.: 6,42e-42 Length: 2103
Score: 475.00 Matches: 97
Percent Similarity: 56.33% Conservative: 41
Best Local Similarity: 39.58% Mismatches: 97
Query Match: 20.28% Indels: 10
DB: 7 Gaps: 5

US-10-803-530-2 (1-435) x US-11-102-240-105 (1-2103)

Qy 190 HisCysLeuAlaCysGly-----LysSerLeu---LysThrProArgValVal 204
Db 557 CATTGC-----TGGCGAACACGAGAGTAAGTAAACTCTAGGTACAGATCTCAGGATCGTT 610
Qy 205 GlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleGlnTrpAspLys 224
Db 611 GGTGGGACAGAAAGTAGAGAGGGTGAATGGCCCTGGCAGGCTAGCTGCAGTGGATGGG 670
Qy 225 GlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAlaAlaHisCys 244
Db 671 AGTCATCGCTGTGGAGCAACCTTAATTAATGACCATGGCTGTGTGAGTCTGCTCACTGT 730
Qy 245 PheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAspLysLeuGly 264

```

```

Db 731 TTTACAACATATATAAGAACCCCTGCCAGATGGACTCTCTTCTTGGAGTAACAATAAACCT 790
Qy 265 SerPheProSerLeuAlaValAlaLysIleIleIleIleGluPheAsnProMetTyrPro 284
Db 791 TCGAAAATGAACGGGTCTCCGAGAAATAATTTCCATGAAATAACAAACACCCATCA 850
Qy 285 LysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSerGlyThrVal 304
Db 851 CATCACTATGATATTTCTCTTGCAGAGCTTTCTAGCCCTGTTCCCTACACAAATGCAGTA 910
Qy 305 ArgProIleCysLeuProPhePheAspGluLeuLeuThrProAlaThrProLeuTrpIle 324
Db 911 CATAGAGTTTGTCTCCCTCATGATCATCTCATGAGTTTCAACGAGGTGATGATGTTGTG 970
Qy 325 IleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeuLeuGlnAla 344
Db 971 ACAGGATTTGGAGCACTGAAA--AATGATGTTTACAGTCAAAATCATCTTCGACAAGCA 1027
Qy 345 SerValGlnValIleAspSerThrArgCysAsnAlaLysAspAlaTyrGlnGlyGluVal 364
Db 1028 CAGGTGACTCTCATAGACGCTACCACTTGCATGAACCTCAAGCTTACAATGACGCCATA 1087
Qy 365 ThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCysGlnGlyAsp 384
Db 1088 ACTCTAGAATGTTATGTGCTGCTCTTAAAGAGAAACAGATGATGCCAGGTGAC 1147
Qy 385 SerGlyGlyProLeuMetTyrGlnSer-----AspGlnTrpHisValValGlyIleVal 402
Db 1148 TCTGGAGACCACTGGTTAGTTTCAGATGCTAGAGATATCTGGTACCTTCTGGAATAGTG 1207
Qy 403 SerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSerAla 422
Db 1208 AGCTGGGAGAGATGAATGTGCGAAACCCCAACGCTGGTGTATTACTAGAGTTACGGCC 1267
Qy 423 TyrLeuAsnTrpIle 427
Db 1268 TTGGGGAGTGGATT 1282

RESULT 6
US-11-137-465-23
; Sequence 23, Application US/11137465
; Publication No. US2005025558A1
; GENERAL INFORMATION:
; APPLICANT: Agarwal, Pankaj
; APPLICANT: Murdoch, Paul R.
; APPLICANT: Rizvi, Safia, K.
; APPLICANT: Smith, Randall, F.
; APPLICANT: Xiang, Zhaoying
; APPLICANT: Kabnick, Karen
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50018
; CURRENT APPLICATION NUMBER: US/11/137,465
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US/10/239,663
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/09226
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 60/192,158
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,668
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/200,166
; PRIOR FILING DATE: 2000-04-27
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 918
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-137-465-23

Alignment Scores:
Pred. No.: 2,79e-42 Length: 918

```

Score: 473.50 Matches: 95
Percent Similarity: 54.98% Conservative: 43
Best Local Similarity: 37.85% Mismatches: 100
Query Match: 20.22% Indels: 13
DB: 7 Gaps: 4

US-10-803-530-2 (1-435) x US-11-137-465-23 (1-918)

```
Qy 193 AlaCysGlyLysSerLeuLysThrProArgValValGlyGluGluAlaSerValAsp 212
Db 94 GCCTGGCGCCACCGGAAATTCACGCGCTGGTGGCGGAGTGAGTGGCGGCGGG 153
Qy 213 SerTrpProTrpGlnValSerIleGlnTrpAspLysGlnHisValCysGlySerIle 232
Db 154 CGCTGGCCATGGCAGGCGAGCTCGCTGGCTGAGGACGCCACCGATGTGGAGGAGCTG 213
Qy 233 LeuAspProHisTrpValLeuThrAlaAlaHisCysPheArgLysHisThrAspValPhe 252
Db 214 CTCAGCCGCGCTGGTGGCTCTCGGCTGGCGACTGCTTCCAAAGACACTACTATCCCTCC 273
Qy 253 AsnTrpLysValArgAlaGlySer-----AspLysLeuGlySer 265
Db 274 GAGTGGAGGTCCAGCTGGCGGAGCTGACTTCCAGGCCAACTCCTTGAACCTGCGGGCC 333
Qy 266 PheProSerLeuAlaValAlaLysIleIleIleIleGluPheAsnProMetTyrProLys 285
Db 334 TACAGCAGTCTTTACAAAGTGCAGGACATATTGTGAACCTGACGCACTTGGGGTTTA 393
Qy 286 AspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSerGlyThrValArg 305
Db 394 CGCAATGACATGTCCTGCTGAGACTGGCTCTTCTGTCACTCAATGCGTACATCCAG 453
Qy 306 ProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrProLeuTrpIleIle 325
Db 454 CCCATTGTCATCGAGTCTTCCACCTTCACTTGTGTGACCGCGGAGTGTGGTGACC 513
Qy 326 GlyTrpGlyPheThrLysGlnAsnGlyCysLysMetSerAsp-----IleLeuLeuGln 343
Db 514 GGCTGGGGTAAATCAGCCCCAGTGGCACACTCTGCCACCTCTTACAACCTCCGGGAA 573
Qy 344 AlaSerValGlnValIleAspSerThrArgCysAsn-----AlaAspAspAlaTyr 360
Db 574 GCACAGTGTACCATTCTAAACACACAGGCTGTAATTACCTGTTTGAACAGCCCTTAGC 633
Qy 361 GlnGlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThr 380
Db 634 CGTAGTAGTATCTGGGATTCATGTTTGTGCTGCTGAGGATGGCAGTGTAGACACC 693
Qy 381 CysGlnGlyAspSerGlyProLeuMetTyrGlnSerAsp---GlnTrpHisValVal 399
Db 694 TGCAAAGGTGACTCAGGTGGACCCCTTGTCTGTGACAAGATGAGTGTGTATCAGGTT 753
Qy 400 GlyIleValSerTrpGlyTyrGlyCysGlyProSerThrProGlyValTyrThrLys 419
Db 754 GGAATCTGTAGCTGGGAAATGAGTGGCGGTCAACCCCAATCGGCTGTGTGTACACCAAC 813
Qy 420 ValSerAlaTyrLeuAsnTrpIleTyrAsnVal 430
Db 814 ATCAGTGTGTACTTCCACTGGATCGGAGGGTG 846
```

RESULT 7

US-10-131-826A-221
; Sequence 221, Application US/10131826A
; Publication No. US20050245730A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330RIC128
; CURRENT APPLICATION NUMBER: US/10/131,826A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 221
; LENGTH: 1129
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-131-826A-221

Alignment Scores:

Pred. No.: 1,74e-41 Length: 1129
Score: 467.50 Matches: 99
Percent Similarity: 55.43% Conservative: 49
Best Local Similarity: 37.08% Mismatches: 92
Query Match: 19.96% Indels: 27
DB: 6 Gaps: 9

US-10-803-530-2 (1-435) x US-10-131-826A-221 (1-1129)

```
Qy 184 GlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrProArgVal 203
Db 81 GGTCTCTCAGAGGGCCAAAGGCGAGCAACAGCCTGTGTGCCCCCAGGATGCTGAACCGAATG 140
Qy 204 ValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleGlnTrpAsp 223
Db 141 GTGGGCGGCGAGGACACGCGAGGAGGCGAGTGGCCCTGGCAAGTCAGCATCCAGCGCAAC 200
Qy 224 LysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAlaAlaHis 243
Db 201 GGAAGCCACTTCTGCGGGGCGAGCCCTCATCGCGAGGAGTGGTCTCTGACGGTGGCGAC 260
Qy 244 CysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAspLysLeu 263
Db 261 TGCTTCGCAACACCTCTGAGACGTCCTGTACAGGTCTGTGGGGGCAAGCACTA 320
Qy 264 ---GlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsnProMet 282
Db 321 GTCCAGCCGGGACCAACAGCTATGTATGCCGGGTGAGGAGGAGGAGCAACCCCTGT 380
Qy 283 Tyr-----ProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThr 299
```

```
Db 381 TACCAGGCGACGGGCTCCAGCGTGCCTGGTGGAGTGGAGCGACCAAGTGGCC 440
Qy 300 PheSerGlyThrValArgProileCysLeuProPhePheAspGluGluLeuThrProAla 319
Db 441 TTCACCAATTACATCCTCCCGTGTGCTGCTCT-----GACCCCTCGGTGATCTTTGAG 494
Qy 320 ThrProLeu-----TrrPilelleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMet 337
Db 495 ACGGCGATGAATGCTGGGTCACTGGCTGGGCGAGCCCGAGTGAG----- 539
Qy 338 SerAsp-----IleLeuLeuGlnAlaSerValGlnValIleAspSer 351
Db 540 GAAGACCTCTCTGCCGAAACCGGATCTCGAGAAACTCGTGTGCCCATCATCGACACA 599
Qy 352 ThrArgCysAsn-----AlaAspAspAlaTyrGlnGlyGlu---Val 364
Db 600 CCCAAGTCGAACCTGCTCTACAGCAAGACACCGAGTTGGCTACCAACCCAAACCATC 659
Qy 365 ThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCysGlnGlyAsp 384
Db 660 AAGATGACATGCTGTGCGCGCTTCGAGGAGGCGCAAGAGGATGCTGCAAGGCGGAC 719
Qy 385 SerGlyGlyProLeuMetTyrGlnSerAspGln---TrpHisValValGlyIleValSer 403
Db 720 TCGGCGCGCCCTGCTGTGCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 779
Qy 404 TrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSerAlaTyr 423
Db 780 TGGGTGAGGGGTGTGCGCGCAGAACCGCCAGGTGTCTACATCCGTGTACCGCGCCAC 839
Qy 424 LeuAsnTrpIleTyrAsnVal 430
Db 840 CACAACCTGATCCATCGGATC 860
```

RESULT 8

```
US-10-821-234-191/c
; Sequence 191, Application US/10821234
; Publication No. US200502551141
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 191
; LENGTH: 1802
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-821-234-191
```

Alignment Scores:

Pred. No.:	3,87e-39	Length:	1802
Score:	449.00	Matches:	113
Percent Similarity:	49.51%	Conservative:	40
Best Local Similarity:	36.57%	Mismatches:	116
Query Match:	19.17%	Indels:	40
DB:	6	Gaps:	10

US-10-803-530-2 (1-435) x US-10-821-234-191 (1-1802)

```
Qy 147 ProThrPheArgAlaValGluIle-----GlyProAspGlnAspLeuAspValValGlu 164
Db 1666 CTAACCTTCGGTGGCTGGAGTGTGCCCGAGGGGCCC----- 1631
Qy 165 IleThrGluAsnSerGlnGluLeuArgMetArgAsnSerGlyProCysLeuSerGly 184
```

```
Db 1630 -----TTGTCTGGGCGCATGGCCAGGAAGGGGGTCTCTGGGCGCTGGCGAGCTGGGG 1580
Qy 185 SerLeuValSerLeuHisCysLeuAla----- 193
Db 1579 GCTGTGGCCATTCTGCTCTATCTTTGGATTACTCCGGTCGGGACAGGAGCGGAAGGGCA 1520
Qy 194 -----CysGlyLysSerLeuLysThrProArgValValGlyGlyGluGluAlaSer 210
Db 1519 GAAGCTCCTCCGCGTGTGGCCCCCAAGCA---CGCATCACAGGTGGCGAGCAGTCAGTC 1463
Qy 211 ValAspSerTrpProTrrPlelleGlnValSerIleGlnTyrAspLysGlnHisValCysGlyGly 230
Db 1462 GCCGGTCAGTGGCCTTGGCAGGTGAGCATCACTATGAAGCGCTCCATGTGTGTGTGGC 1403
Qy 231 SerIleLeuAspProHisTrpValLeuThrAlaAlaHisCysPheArgLysHisThrAsp 250
Db 1402 TCTCTCGTGTGAGCAGTGGGTGCTGTCAGCTGCTCACTCTTCCCGCAGCGAGCACCAC 1343
Qy 251 ValPheAsnTrpLysValArgAlaGlySerAspLysLeuGlySerPheProSerLeuAla 270
Db 1342 AAGGAAGCCTATGAGTCAAGCTGGGGGCCCAACAGCAGCTAGACTCTCTACTCCGAGGACGCC 1283
Qy 271 ---ValAlaLysIlellellellellellellellellellellellellellellelle 286
Db 1282 AAGGTCAAGCACCCCTGAAGGACATCATCCCCCAGCTACCTCCAGGAGGGCTCCCCAG 1223
Qy 287 AsnAspIleAlaLeuMetLeuGlnPheProLeuThrPheSerGlyThrValArgPro 306
Db 1222 GCGGACATGACACTCTCCAACTCAGAGACCCATCACTTCTCCCGCTACATCCGGCCC 1163
Qy 307 IleCysLeuProPhePheAspGluGluLeuThrProAlaThrProLeuTrpIlellelle 326
Db 1162 ATCTGCTCTCCCTGCAGCCACGCCCTCTTCCCCAAGCGCTCCACTGCACCTGCTACTGCG 1103
Qy 327 TrpGlyPheThrLysGlnAsnGlyGlyLysMetSer---AspIleleuLeuGlnAlaSer 345
Db 1102 TGGGTCATGTGGCCCCCTCAGTGAGCTCTCTGAGCGCCCAAGCCACTGCAGCAACTCGAG 1043
Qy 346 ValGlnValIleAspSerThrArgCysAsnAla-----AspAspAlaTyrGlnGly 362
Db 1042 GTGCTCTGATCAGTCGTCGAGACGTCGTAACTGCTGTACAAATCGAGCCCAAGCCTGAG 983
Qy 363 Glu-----ValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAsp 379
Db 982 GAGCGCGCACTTGTCCAGAGACATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 923
Qy 380 ThrCysGlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAsp---GlnTrpHisVal 398
Db 922 GCCTGCCAGGGTGACTCTGGGGGCCCACTCTCTGCTGCTGTGGAGGGTCTCTGTGTACCTG 863
Qy 399 ValGlyIleValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThr 418
Db 862 ACGGGCATTTGAGCTGGGGAGATGCCTGTGGGGGCCGCAACAGGCGCTGTGTGTACACT 803
Qy 419 LysValSerAlaTyrLeuAsnTrpIle 427
Db 802 CTGGCTCCAGCTATGCTCTCTGGATC 776
```

RESULT 9

```
US-11-056-621-5
; Sequence 5, Application US/11056621
; Publication No. US20050262592A1
; GENERAL INFORMATION:
; APPLICANT: Spencer, David
; APPLICANT: Dickey, Lynn F.
; APPLICANT: Gasdaska, John R.
; APPLICANT: Wang, Xiaowei
; APPLICANT: Cox, Kevin M.
; APPLICANT: Peele, Charles G.
; TITLE OF INVENTION: EXPRESSION OF PLASMINOGEN AND
; TITLE OF INVENTION: MICROPLASMINOGEN IN DUCKWEED
; FILE REFERENCE: 40989/274646
```

```

; CURRENT APPLICATION NUMBER: US/11/056,621
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: US 60/543,487
; PRIOR FILING DATE: 2004-02-11
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 780
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Duckweed codon optimized sequence encoding human
; OTHER INFORMATION: microplasma
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(780)
US-11-056-621-5

Alignment Scores:
Pred. No.: 5,938-39 Length: 780
Score: 442.50 Matches: 97
Percent Similarity: 52.61% Conservative: 44
Best Local Similarity: 36.19% Mismatches: 98
Query Match: 18.89% Indels: 29
DB: 7 Gaps: 7

US-10-803-530-2 (1-435) x US-11-056-621-5 (1-780)
QY 176 AnSerSerGlyProCysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGly 195
Db 22 AACACGACAGTGGTGAACCGAAGAGTCCCTGGCCGCGTGGAGGCGTGGTGGCCAC 111
QY 196 LysSerLeuLysThrPro-----ArgValValGlyGluGluAlaSer 210
Db 52 AAGCCACAGTGGTGAACCGAAGAGTCCCTGGCCGCGTGGAGGCGTGGTGGCCAC 111
QY 211 ValAspSerTrpProTrpGlnValSerLeuGlnTyrAsp---LysGlnHisValCysGly 229
Db 112 CCGACTCTCTGCGCCCTGGCAAGTACGCTGGCCGACCCGCTTCGGCATGCACTTCTGGGC 171
QY 230 GlySerIleLeuAspProHisTrpValLeuThrAlaHisCysPheArgLysHisThr 249
Db 172 GGCACCTCATCTCCCGAGTGGTGTCTGACCGCGCTCACTGCTCGAAGAGTCCCG 231
QY 250 AspValPheAsnTrpLysValArgAlaGlySerAspLysLeuGlySerPhe----- 266
Db 232 AGGCCCTCTCTACAGGTATCTCTGGCGCCGCCACGAGGAGTGAACCTCGAGCCGCCAC 291
QY 267 ---ProSerLeuAlaValAlaLysIleIleIleIleIleGluPheAsnProMetTrpProLys 285
Db 292 GTTCAGGAGATCGAGGTGTCCCGTGTTC-----CTGGAGCCGCCAG 333
QY 286 AspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSerGlyThrValArg 305
Db 334 CGCAAGGATATCGCCCTGCTCAAGCTCTTAGCCCGCGCGTCACTACCGACAGGTTATC 393
QY 306 ProIleCysLeuProPhePheAspGluLeuThrProAlaThrProLeuTrpIleIle 325
Db 394 CCGCGCTGCTTCCCTCCCGAATACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 453
QY 326 GlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeuLeuGlnAlaSer 345
Db 454 GCGTGGGCGAGACCCAGGAAACGTTCCGCG-----CGCGGCTCTCTAAGGAGGCCAG 507
QY 346 ValGlnValIleAspSerThrArgCysAsnAlaAspAlaThrGlnGlyGluValThr 365
Db 508 CTCCCGGTGATTGAGACACAGGTGTGCAACGGTTACGAGTTCCTGAACCGCGCGTCCAG 567
QY 366 GluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCysGlnGlyAspSer 385
Db 568 TCACCGCAACTCTGCGCGGCGACTTGGCGCGGCGGACGACGACGACGACGACGACGACG 627
QY 386 GlyGlyProLeuMet---TyrGlnSerAspGlnTrpHisValValGlyIleValSerTrp 404

```

```

Db 628 GCGCGGCGCTGGTGTCTCGAAGAGACAAAGTACATCTCTCAAGGCGTCACTGCTGG 687
QY 405 GlyTyrGlyCysGlyGlyProSerThrProGlyValThrLysValSerAlaTyrLeu 424
Db 688 GGCCTCGCTGCGCACGCCCTAACAGCGGCGCTCTATGTGCGCGTGTCCCGCTCGTG 747
QY 425 AsnTrpIleTyrAsnValTrpLys 432
Db 748 ACCTGGATCGAGGCGTGTATGCGG 771

RESULT 10
US-11-056-621-3
; Sequence 3, Application US/11056621
; Publication No. US20050262592A1
; GENERAL INFORMATION:
; APPLICANT: Spencer, David
; APPLICANT: Dickey, Lynn F.
; APPLICANT: Gasdaska, John R.
; APPLICANT: Wang, Xiaowei
; APPLICANT: Cox, Kevin M.
; APPLICANT: Peele, Charles G.
; TITLE OF INVENTION: EXPRESSION OF PLASMINOGEN AND
; FILE REFERENCE: 40989/274646
; CURRENT APPLICATION NUMBER: US/11/056,621
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: US 60/543,487
; PRIOR FILING DATE: 2004-02-11
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2376
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Duckweed codon optimized sequence encoding human
; OTHER INFORMATION: plasminogen
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2376)
US-11-056-621-3

Alignment Scores:
Pred. No.: 3,928-38 Length: 2376
Score: 441.50 Matches: 114
Percent Similarity: 48.00% Conservative: 54
Best Local Similarity: 32.57% Mismatches: 133
Query Match: 18.85% Indels: 49
DB: 7 Gaps: 10

US-10-803-530-2 (1-435) x US-11-056-621-3 (1-2376)
QY 124 AlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMetGlyTyr 143
Db 1346 TCTCTGCTCCGACGCTCGAGACCCCGTCCGAGGAGGACTGTTCGGGAACGGAAGG 1405
QY 144 SerSerLysProThrPheArgAlaValGluIleGlyProAsp---GlnAspLeuAspVal 163
Db 1406 GCTACCGCGGCAAGCGCGACCACTGTACCGGACCCCGTCCGAGGACTTGGCGCGCTC 1465
QY 163 alGluIleThrGluAsnSer-----GlnG 171
Db 1466 AGGAGCCCGACGAGCACAGCATTTTACCCCGAGACCAACCCGCGCGCGCTGGAGA 1525
QY 171 luLeuArgMetArgAsnSerSerGlyPro-----CysLeuSerGlySerL 186
Db 1526 AGAATACTGCGGCAACCCGACGCGGACGCTGTGGAGGCCCTGTGTACTACTACCAAC 1585
QY 186 euValSerLeuHis-----CysLeuAla- 195
Db 1586 CGCGTAAGCTCTACGACTACTGTGCGAGCTGCGAGTGGCGGCTCCGCTCTTCGACTGG 1645

```

```
QY 195 lyLysSerLeuLysThrPro-----ArgValValGlyGlyGluGluAlas 210
Db 1646 GGAAGCCACAGGTGGAACCGAAGAAGTGCCTCGCGCGGTGTGTGGAGGGTGGTGGCGCC 1705
QY 210 erValAspSerTrpProTrpGlnValSerIleGlnTyrAsp---LysGlnHisValCysG 229
Db 1706 ACCCGCACTCTCGCGCCCTGGCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 1765
QY 229 lyGlySerIleLeuAspProHisTrpValLeuThrAlaAlaHisCysPheArgLysHisIst 249
Db 1766 CGCGACCTCATCTCCCGGAGTGGTTCAGCCGCGCTCACTGCTCGAGAGTCC 1825
QY 249 hrAspValPheAsnTrpLysValArgAlaGlySerAspLysLeuGlySerPhe----- 266
Db 1826 CGAGGCCCTCTCTACAAAGTTCATCTCGGGGCCACACAGGAGTGAACCTCGAGCGCC 1885
QY 267 -----ProSerLeuAlaValAlaLysIleIleIleGluPheAsnProMetTyrProL 285
Db 1886 AGTTTCAGGAGATCGAGGTGTCCCGCTTGTTC-----CTGGAGCCCA 1927
QY 285 ysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSerGlyThrVala 305
Db 1928 CGCGCAAGATATCCCTCTCAAGTCTCTAGCCCGCGCGTCAATCACCGAAGGTGA 1987
QY 305 rgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrProLeuTrpIleI 325
Db 1988 TCCCGCGCTCTCTCCCGCAACTAGTGTGCTGACCGCACCGAGTCTTCGTGA 2047
QY 325 leGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeuLeuGlnAlas 345
Db 2048 CGCGTGGCGGAGACCCAGGAACTTGGC-----GCGGCGCTCTCAAGGAGGCC 2101
QY 345 erValGlnValIleAspSerThrArgCysAsnAlaAspAlaTyrGlnGlyGluValT 365
Db 2102 AGCTCCCGTGATTGAGAACAGAGTGTGCACCGTTACAGTTCCTCAACCGCGCGTCC 2161
QY 365 hrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCysGlnGlyAspS 385
Db 2162 AGTCACCGCAACTCTGCGCGCGGCACTTGGCGCGCGGACCCAGACAGTCCAGGCGGACA 2221
QY 385 erGlyGlyProLeuMet---TyrGlnSerAspGlnTrpHisValValGlyIleValSerT 404
Db 2222 CGCGCGCGCGCTGTGTGTCTTCGAGAGGACAAGTACATCTCCAGCGCGTCAAGTCTCT 2281
QY 404 rpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSerAlaTyrL 424
Db 2282 GGGGCTCGGCTGGCAGCCCTAACAGCGCGGCTCTATGTGCGCGTGTCCCGCTTCG 2341
QY 424 euAsnTrpIleTyrAsnValTrpLys 432
Db 2342 TGACCTGGATCGAGGCGGTGATGCGC 2367
```

RESULT 11

```
US-10-131-826A-505
; Sequence 505, Application US/10131826A
; Publication No. US20050245730A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tamas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
```

```

; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C128
; CURRENT APPLICATION NUMBER: US/10/131,826A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 505
; LENGTH: 1204
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-131-826A-505

Alignment Scores:
Pred. No.: 1,76e-31 Length: 1204
Score: 377.50 Matches: 99
Percent Similarity: 49.15% Conservative: 45
Best Local Similarity: 33.79% Mismatches: 111
Query Match: 16.12% Indels: 38
DB: 6 Gaps: 11

US-10-803-530-2 (1-435) x US-10-131-826A-505 (1-1204)
QY 163 ValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSer----- 178
Db 18 GTTAGGTGGCTGCGGGACTGGAAGTCACTCGGGCAGAGGTCTCACAGCAGCCAAGGAACC 77
QY 179 ---GlyProCysLeu-----SerGlySerLeuValSerLeuHisCys-Le 192
Db 78 TGGGGCGCGCTCTCTCCCGCTCCAGGCCATGAGGATTCTGCAGTTAATCTGCTGCTCT 137
QY 192 uAlaCysGlyLysSerLeuLysThrProArgValValGlyGlyGluAlaSerValas 212
Db 138 GGCACACAGGCTTGTAGGGGAGAGACCAGGATCATCAAGGGGTTCGAGTGCAGGCTCA 197
QY 212 pSerTrpProTrpGlnValSerIleGlnTyrAspLysGlnHisValCysGlySerI 232
Db 198 CTCCAGCGCTGGCAGGAGCCCTGTTTCGAGAGAGCGCGGTACTCTGTGGGCGGCGCT 257
QY 232 eLeuAspProHisTrpValLeuThrAlaAlaHisCysPheArgLys-----Hi 248
Db 258 CATCCCGCCCGCATGGCTCTCGACAGCAGCCCACTGCCTCAAGCCCGCTACATAGTTCA 317
QY 248 sThrAspValPheAsnTrpLysValArgAlaGlySerAspLys-----LeuGl 264
Db 318 CTGGGGCAGCACAACCTCCAGAGGAGGAGGCGCTGTGAGCAGACCCGCGACGCACTGA 377
QY 264 ySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsnProMetTyrPr 284
Db 378 GTCTCTCCCGCCCGCGGC-----TTCAACAACAGCCTCC 413
QY 284 o---LysAsp-----AsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSe 301
```

```
Db 414 CAACAAAGACCAACCCGCAATGACATCATGCTGTGAAGATGGCATCGCCAGTCTCCATCAC 473
QY 301 rGlyThrValArgProIleCysLeuProPheAspGluLeuThrProAlaThrPr 321
Db 474 CTGGGCTGTGCACCCCTCACCTCTCC-----TCACGCTGTGTCACTGTGGCACCAG 527
QY 321 oLeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLe 341
Db 528 CTGCTCATTTCCGGCTGGGGCAGCAGCTGCAGCCCCCAGTTACCGCTGCCTCACACCTT 587
QY 341 uLeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaTyrG1 361
Db 598 GCGATGCCCAACATCACCATCATTTGAGCACCAGAGTGT-----GAGAACGCTACCC 641
QY 361 nGlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCy 381
Db 642 CGGCAACATCACAGACCATGTGTGTGCCAGCGTGCAGGAAGGGGCAAGGACTCTCTG 701
QY 381 sGlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyI1 401
Db 702 CCAGGCTGACTCCGGGGCCCTCTGTCTGTAAACCATGCT-----CTTCAAGGCAT 752
QY 401 eValSerTrpGlyTyrGly---CysGlyGlyProSerThrProGlyValTyrThrLysVa 420
Db 753 TATCTCTGGGGCCAGGATCCGTGTGCCATCACCCGAAAGCCTGGTGTCTACAGAAAGT 812
QY 420 lSerAlaTyrLeuAsnTrpIleTyrAsnValTrpLys 432
Db 813 CTGCAAAATATGTGACTGGATCCAGGACGATGAAG 849
```

RESULT 12

```
US-11-029-003-7
; Sequence 7, Application US/11029003
; Publication No. US20050260194A1
; GENERAL INFORMATION:
; APPLICANT: PETERS, ROBERT T.
; APPLICANT: MEZO, ADAM R.
; APPLICANT: RIVERA, DANIEL S.
; APPLICANT: BITONTI, ALAN J.
; APPLICANT: STATTEL, JAMES
; TITLE OF INVENTION: IMMUNOGLOBULIN CHIMERIC MONOMER-DIMER HYBRIDS
; FILE REFERENCE: 08945.0007-01000
; CURRENT APPLICATION NUMBER: US/11/029,003
; CURRENT FILING DATE: 2005-01-05
; PRIOR APPLICATION NUMBER: 60/539,207
; PRIOR FILING DATE: 2004-01-26
; PRIOR APPLICATION NUMBER: 60/487,964
; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 60/469,600
; PRIOR FILING DATE: 2003-05-06
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 7
; LENGTH: 2016
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: construct
US-11-029-003-7
```

```
Alignment Scores:
Pred. No.: 2,88e-30 Length: 2016
Score: 369.50 Matches: 118
Percent Similarity: 41.65% Conservative: 54
Best Local Similarity: 28.57% Mismatches: 127
Query Match: 15.78% Indels: 114
DB: 7 Gaps: 16
```

US-10-803-530-2 (1-435) x US-11-029-003-7 (1-2016)

QY 84 GlyGluAspGlu-----GluHisCysValLysSer 93

```
Db 216 GGAGAGACCAAGCTGTTCTGGATTCTTACAGTGTGGGACCAAGTGTGCTCAAGTCC 275
QY 94 PheProGlu-----GlyProAlaValAlaValArgLeu----- 104
Db 276 ATGCCAGAATGGGGCTCTGCAAGGACCAAGCTCCAGTCTCTATATCTGCTTCTGCTCTCC 335
QY 105 -----SerLys 106
Db 336 TGCTTCGAGGGCGGGAACCTGTGAGAGCACAAGGATGACCAAGCTGATCTGTGTGAACGA 395
QY 107 AspArgSerThrLeuGlnValLeu-AspSerAlaThrGlyAsnTrpPheSerAla---Cy 125
Db 396 GAACGGCGGCTGTGAGCAGTACTGACGAGTACCCACGCGGACCAAGCGCTCTGTCTCGGTG 455
QY 125 sPheAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMetGlyTyrSerSe 145
Db 456 CCACGAGGGGTACTCT---CTGCTGGCAGAC-----GGGGTGTCTGTG 494
QY 145 rLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAspValValGluI1 165
Db 495 CACACCCACA-----GTTGAATAT---CCATGTGGAAAAAATACCTATTCTAGAA-- 540
QY 165 eThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCysLeuSerGlySe 185
Db 541 -----AAAAGAAATGCCAGCAAAACCC----- 561
QY 185 rLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrProArgValValG1 205
Db 562 -----CAAGGCGCAATTGTGGG 578
QY 205 yGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleGlnTyrAspLysG1 225
Db 579 GGGCAAGGTGTGCCCAAAAGGGAGTGTCCATGCCAGGTCTCTGTGTGTGTAATGGAGC 638
QY 225 nHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAlaAlaHisCysPh 245
Db 639 TCAGTTGTGTGGGGGAGCCCTGATCAACACCATCTGGTGTGTCTCGCGGCCCACTGTTT 698
QY 245 eArgLysHisThrAspValPheAsnTrpLysValArgAlaGly----- 259
Db 699 CGACAAA-----ATCAAGAACTGGAGAACTGATCGCGGTGCTGGCGGACGACGA 749
QY 260 ----SerAspLysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleG1 278
Db 750 CCTCAGCGACGACGACGGGATGAGCAGAGCGCGGTGGCGAGTCAATCATCCCCAG 809
QY 278 uPheAsnProMetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLe 298
Db 810 CACGTACGTCCCGGGCACCACCAACACGACATCGCGCTGCTCCGCTGCACACGCGCGT 869
QY 298 uThrPheSerGlyThrValArgProIleCysLeuPro-----PhePheAspGluG1 315
Db 870 GGTCTCCTCAGTACCATGTGTGTCCTCTGCTCCCGAACGAGCATTTCTCTGAGGAGAC 929
QY 315 uLeuThrProAlaThrProLeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyG1 335
Db 930 GCTGGCTTCTGTGCGCTTCTCATTTGTCAGCGSCTGGGGCCAGCTGCTGGACCGTGGCGC 999
QY 335 yLysMetSerAspIleLeu-----LeuG1 343
Db 990 CACGGCCCTGGAGCTCATGGTCTCTCAACGTGCCCGGCTGATGACCCAGGACTGCTGCA 1049
QY 343 nAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaTyrGlnGlyG1 363
Db 1050 GCAGTCACGAAGGTGGGAGACTCCCA-----AA 1079
QY 363 uValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCysGlnG1 383
Db 1080 TATCAGGAGTACATGTTCTGTGCGCGTACTCGGATGGCAGCAAGACTCTCTGCAAGGG 1139
QY 383 yAspSerGlyGlyPro---LeuMetTyrGlnSerAspGlnTrpHisValValGlyIleVa 402
```



```
Db 1140 GGACAGTGGAGGCCACACCTACCGGGGACGTGGTACCTGACGGGCATCGT 1199
Qy 402 lserTrpGlyTyGlyCysGlyGlyProSerThrProGlyValTyThrLysValSerAl 422
Db 1200 CAGCTGGGGCCAGGCGTGGCAACCGTGGGCCACCTTTGGGGTGTACACAGGGGTCTCCCA 1259
Qy 422 atYrLeuAenTrpIleTyAenValTrpLysAlaGlu 434
Db 1260 GTATATCGAGTGGCTGGCAAGAGCTCATGGGCTCAGAG 1296

RESULT 13
US-10-623-155-122
; Sequence 122, Application US/10623155
; Publication No. US20050261166A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Peckham, David W.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C20
; CURRENT FILING DATE: US/10/623,155
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 122
; LENGTH: 1475
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-623-155-122

Alignment Scores:
Pred. No.: 5,686-29 Length: 1475
Score: 356.00 Matches: 91
Percent Similarity: 49.25% Conservative: 40
Best Local Similarity: 34.21% Mismatches: 103
Query Match: 15.20% Indels: 32
DB: 6 Gaps: 9

US-10-803-530-2 (1-435) x US-10-623-155-122 (1-1475)
Qy 194 CysGlyLysSerLeuLysThrProArg-----ValValGlyGlyGluAlaSerVal 211
Db 582 TGTGGCCAAAGAGACTCTGAGGCGCCGCTTTAAGATTATTGGGGGAGAATTCACCAACATC 641
Qy 212 AspSerTrpTrpGlnValSerIleGlnTyAspLysGln----- 225
Db 642 GAGAACCCAGCCCTGGTTTGGCGCCATC---TACAGGAGGACCGGGGGGCTGTGTACC 698
Qy 226 HisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAlaAlaHisCysPhe 245
Db 699 TACGTGTGTGGAGGAGCGCTCATCAGCCCTTCTGGGTGATCAGCGCCACACACTGCTTC 758
Qy 246 ArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAspLysLeuGlySer 265
Db 759 ATTGATTACCCAAAGAGGAGACTACATCGTCTACCTGGGTGGTCAAGGCTTAACCTCC 818
Qy 266 PhePro-----SerLeuAlaValAlaLysIleIleIleIleIleGluPheAsnPro 281
Db 819 AACACGCAAGGGGAGATGAAGTTTGAGGTGGAAACCTCATCCTA-----CACAG 869
Qy 282 MetTyProLysAsp-----AsnAspIleAlaLeuMetLysLeuGlnPhe 296
Db 870 GACTACAGCGCTGACACGCTTGTCTACCAACAGCATTTGCCTTCTGTAAGATCCGCTCC 929
Qy 297 -----ProLeuThrPheSerGlyThrValArgProIleGlyCysLeu 309
Db 930 AAGGAGGGCAGGTGTGCGCAGCCA-----TCCCGGACTATACAGACCATCTGCTG 980
Qy 310 ProPhePheAspGluGluLeuThrProAlaThrProLeuTrpIleIleGlyTrpGlyPhe 329
Db 981 CCCTCGATGTATAACGATCCCGAGTGTGGCACAAAGCTGTGAGATCACTGGCTTTGGAAA 1040
```

```
Qy 330 ThrLysGlnAenGlyGlyLysMetSerAspIleLeuLeuGlnAlaSerValGlnValIle 349
Db 1041 GAGAAATCTACCGACTATCTCTATCCGAGCAGCTGAAAGATGACTGTTGTGAAGCTGATT 1100
Qy 350 AspSerThrArgCysAsnAlaAspAlaTyTrpGlnGlyGluValThrGluLysMetMet 369
Db 1101 TCCACACGGGAGTGTACGACGCCCTACTACGGCTCTGAAGTCCACCAAAATGCTG 1160
Qy 370 CysAlaGlyIleProGluGlyValAspThrCysGlnGlyAspSerGlyGlyProLeu 389
Db 1161 TGTGCTGTGACCCACACATGGAAAAACATCTCTCCAGGGAGACTCAGGGGGACCCCTC 1220
Qy 390 MetTyGln---SerAspGlnTrpHisValValGlyIleValSerTrpGlyTyGlyCys 408
Db 1221 GTCTGTTCCTCCCAAGCGCGCATGCTTGTGAGTTGTGAGTGGGGCGGTGGATGT 1280
Qy 409 GlyGlyProSerThrProGlyValTyThrLysValSerAlaTyLeuAenTrpIleTy 428
Db 1281 GCCCTGAAGGACAGCCAGGCGTCTACACGAGAGTCTCACACTTCTTACCTTGGATCCG 1340
Qy 429 AsnValTrpLysAlaGlu 434
Db 1341 AGTCACACCAAGGAAGAG 1358

RESULT 14
US-10-623-155-123
; Sequence 123, Application US/10623155
; Publication No. US20050261166A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Peckham, David W.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C20
; CURRENT FILING DATE: US/10/623,155
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 123
; LENGTH: 2294
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-623-155-123

Alignment Scores:
Pred. No.: 1,098-28 Length: 2294
Score: 356.00 Matches: 91
Percent Similarity: 49.25% Conservative: 40
Best Local Similarity: 34.21% Mismatches: 103
Query Match: 15.20% Indels: 32
DB: 6 Gaps: 9

US-10-803-530-2 (1-435) x US-10-623-155-123 (1-2294)
Qy 194 CysGlyLysSerLeuLysThrProArg-----ValValGlyGlyGluAlaSerVal 211
Db 582 TGTGGCCAAAGAGACTCTGAGGCGCCGCTTTAAGATTATTGGGGGAGAATTCACCAACATC 627
Qy 212 AspSerTrpTrpGlnValSerIleGlnTyAspLysGln----- 225
Db 628 GAGAACCCAGCCCTGGTTTGGCGCCATC---TACAGGAGGACCGGGGGGCTGTGTACC 684
Qy 226 HisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAlaAlaHisCysPhe 245
Db 685 TACGTGTGTGGAGGAGCGCTCATCAGCCCTTCTGGGTGATCAGCGCCACACACTGCTTC 744
Qy 246 ArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAspLysLeuGlySer 265
Db 745 ATTGATTACCCAAAGAGGAGACTACATCGTCTACCTGGGTGGTCTCAAGGCTTAACCTCC 804
```

266	PhePro-----	SerLeuAlaValAlaLysIleIleIleIleGluPheAsnPro	281
805	AACACGCAAGGGGAGATGAAGTTTGAGGTGGAAACCTAATCTTA-----	CAACA	855
282	MetTyrProLysAsp-----	AsnAspIleAlaLeuMetLysLeuGlnPhe	296
856	GACTACAGCGCTGACACGCTTGCTCACCACAACGACATTGCCTTGCTGAGATCCGTTCC	915	
297	-----	ProLeuThrPheSerGlyThrValArgProIleCysLeu	309
916	AAGGAGGGCAGGTGTCGCAGCCA-----	TCCCGACTATACAGACCATCTGCCTG	966
310	ProPhePheAspGluLeuThrProAlaThrProLeuTrpIleIleGlyTrpGlyPhe	329	
967	CCCTCGATGATAACGATCCCGAGTTTGGCACAGCTGTGAGATCACTGGCTTTGGAAA	1026	
330	ThrLysGlnAsnGlyLysMetSerAspIleLeuGlnAlaSerValGlnValIle	349	
1027	GAGAAATTACCGACTATCTCTATCCGAGCAGCTGAAATGACTGTTGTGAAGCTGATT	1086	
350	AspSerThrArgCysAsnAlaAspAspAlaTyrGlnGlyGluValThrGluLysMetMet	369	
1087	TCCCACCGGAGTGTACGACGCCCACTACTACGGCTCTGAAGTCACCAACAAATGCTG	1146	
370	CysAlaGlyIleProGluGlyValAspThrCysGlnGlyAspSerGlyGlyProLeu	389	
1147	TGTGTGTCGACCACACAGATTCCTGCGCAGGGAGACTCAGGGGGACCCCTC	1206	
390	MetTyrGln---SerAspGlnTrpHisValValGlyIleValSerTrpGlyTyrGlyCys	408	
1207	GTCGTCCCTCCCAAGCGCCGATGACTTTGACTGGAAATTGTGAGCTGGGGCCGTGGATGT	1266	
409	GlyGlyProSerThrProGlyValTyrThrLysValSerAlaTyrLeuAsnTrpIleTyr	428	
1267	GCCCTGAAGGACAGCCAGCGTCTACAGAGAGTCTCACACTTCTTACCCGTGATCCGC	1326	
429	AsnValTrpLysAlaGlu	434	
1327	AGTCACACCAAGGAAG	1344	

```

RESULT 15
US-11-029-003-9
; Sequence 9, Application US/11029003
; Publication No US20050260194A1
; GENERAL INFORMATION:
; APPLICANT: PETERS, ROBERT T.
; APPLICANT: MEZO, ADAM R.
; APPLICANT: RIVERA, DANIEL S.
; APPLICANT: BITONTI, ALAN J.
; APPLICANT: STATTEL, JAMES
; TITLE OF INVENTION: IMMUNOGLOBULIN CHIMERIC MONOMER-DIMER HYBRIDS
; FILE REFERENCE: 08945.0007-01000
; CURRENT APPLICATION NUMBER: US/11/029,003
; CURRENT FILING DATE: 2005-01-05
; PRIOR APPLICATION NUMBER: 60/539,207
; PRIOR FILING DATE: 2004-01-26
; PRIOR APPLICATION NUMBER: 60/487,964
; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 60/469,600
; PRIOR FILING DATE: 2003-05-06
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 9
; LENGTH: 2091
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: construct
US-11-029-003-9

```

Alignment Scores:	
Pred. No.:	1.08e-28
Length:	2091

Score:	355.50	Matches:	109
Percent Similarity:	41.8%	Conservative:	50
Best Local Similarity:	28.6%	Mismatches:	144
Query Match:	15.1%	Indels:	77
DB:	7	Gaps:	15

US-10-803-530-2 (1-435) x US-11-029-003-9 (1-2091)

QY	81	CysProLeuGluAspGluHisCysValLysSerPheProGluGlyProAlaVal	100
DB	355	TGTCCTTTGGATTGGAAGAAAGAACTGTGAATTAGAT-----GTA	396
QY	101	AlaValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsn	120
DB	397	ACATGTAAACATTAAAGATGCAGATCGAGCAGCTTTTGTAAAAATAGTCTGTATAACAAG	456
QY	121	TrpPheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArg---	139
DB	457	GTGGTTTGCTCTCTACTGAGGATATGACGTTCAGAAAAACAGAAAGTCCTGTGAACCA	516
QY	140	-----GlnMetGlyTyrSerSerLysProThrPheArg	150
DB	517	GCAGTGCCATTTCCATGTGGAAGAGTTTCTGTGTTTCAAACTTCTAAGCTCACC---	573
QY	151	AlaValGluIleGlyProAspGlnAsp	160
DB	574	GCTGAGACTGTTTTTCCGTGATGTGACTATGTAAATTTCTACTGAAGCTGAAACCAATT	633
QY	161	AspValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyPro	180
DB	634	GAT-----AACATCACTCAAGACCAACCATTTAAT-----	666
QY	181	CysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThr	200
DB	667	-----GACTTC	672
QY	201	ProArgValValGlyGluGluAlaSerValAspSerTrpTrpGlnValSerIle	220
DB	673	ACTCGGGTGTGTGGAGAAGATGCCAAACAGGTCAATTCCTCTGCAGTGTGTTTG	732
QY	221	GlnTyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThr	240
DB	733	AATGTTAAAGTTGATGCATTCCTGGAGGCTCTATCGTTAATGAAATGAATGTGAAC	792
QY	241	AlaAlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySer	260
DB	793	GCTGCCACATGTGTGAAACATGGTGTTTAAAT-----ACAGTTGTGCAGAGTGAA	843
QY	261	AspLysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIle	276
DB	844	CATAATATTGAGGAGACAGACATACAGACCAAGCGAAATGTGATTGCAATATTTCCT	903
QY	277	---IleGluPheAsnProMetTyrProLysAspAsn---AspIleAlaLeuMetLysLeu	294
DB	904	CACCACAACATACAAATGACGCTATTAAATAAGTACAACCATGACATTCCTTCGGAAC	963
QY	295	GlnPheProLeuThrPheSerGlyThrValArgTrpIleCysLeuProPhePheAspGlu	314
DB	964	GACGAACCCTTAGTGTAACACGCTACGTTTACACCTATTGTTCATTCCT-----GACAAG	1017
QY	315	GluLeuThr-----ProAlaThrProLeuTrpIleIleGlyTrpGly	328
DB	1018	GAATACAGCAACATCTTCCTCAAAATTTGGATCTGGCTATGTAAAGTGGCTGGGAAGAGTC	1077
QY	329	PheThrLysGlnAsnGlyGlyLysMetSerAspIleLeuLeuGlnAlaSerValGlnVal	348
DB	1078	TTCCACAAA-----GGGAGATCAGCTTTTAGTTCITTCAGTACCTAGAGTTCCACT	1128
QY	349	IleAspSerThrArgCysAsnAlaAspAspAlaTyrGlnGluValThrGluLysMet	368
DB	1129	GTTCACCGGACCATGCTTCGTATCATCAAGTTC-----ACCATTATAACAAACATG	1182
QY	369	MetCysAlaGlyIleProGluGlyValAspThrCysGlnGlyAspSerGlyGlyPro	388

R:Keime, S.; Adham, I.M.; Engel, W.
Eur. J. Biochem. 190, 195-200, 1990
A>Title: Nucleotide sequence and exon-intron organization of the human proacrosin gene.
A:Reference number: S11674; MUID:90306003; PMID:2114285
A:Accession: S11674
A:Molecule type: DNA
A:Residues: 1-421 <KE1>
A:Cross-references: UNIPROT:P10323; UNIPARC:UPI000012539D; EMBL:X54017; NID:q35582; PIDN
A>Note: the authors translated the codon AGG for residue 64 as Thr and CTG for residue 2
R:Vaquez-Levin, M.H.; Revencos, J.; Gordon, J.W.
Eur. J. Biochem. 207, 23-26, 1992
A>Title: Molecular cloning, sequencing and restriction mapping of the genomic sequence
A:Reference number: S23499; MUID:92331659; PMID:1628652
A:Accession: S23499
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-421 <VA2>
A:Cross-references: UNIPARC:UPI000012539D; EMBL:M77378
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1992
R:Keime, S.
submitted to the EMBL Data Library, December 1989
A:Reference number: S12063
A:Accession: S12063
A:Molecule type: DNA
A:Residues: 1-225, 'R', 227-421 <KE12>
A:Cross-references: UNIPARC:UPI0000172B0D; EMBL:X54017
R:Adham, I.M.; Klemm, U.; Water, W.M.; Engel, W.
Hum. Genet. 84, 125-128, 1990
A>Title: Molecular cloning of human preproacrosin cDNA.
A:Reference number: A61022; MUID:90128988; PMID:2298447
A:Accession: A61022
A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-63, 'T', 65-225, 'V', 227-267, 'R', 269-421 <ADH>
A:Cross-references: UNIPARC:UPI0000172B0E
R: Baba, T.; Watanabe, K.; Kashiwabara, S.I.; Arai, Y.
FEBS Lett. 244, 296-300, 1989
A>Title: Primary structure of human proacrosin deduced from its cDNA sequence.
A:Reference number: S03350; MUID:89153568; PMID:2493394
A:Accession: S03350
A:Molecule type: mRNA
A:Residues: 1-63, 'T', 65-119, 'V', 121-165, 'U', 167-267, 'R', 269-344, 'R', 346-421 <BAB>
A:Cross-references: UNIPARC:UPI0000167BCA; EMBL:Y00970; NID:q28325; PIDN:CAA6784.1; PID
C:Genetics:
A:Gene: GDB:ACR
A:Cross-references: GDB:119645; OMIM:102480
A:Map position: 22q13-22qter
A:Introns: 26/2; 94/2; 189/2; 237/3
C:Keywords: glycoprotein; hydrolase; serine proteinase; sperm
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-421/Product: acrosin #status predicted <MAT>
F:43-421/Product: acrosin heavy chain #status predicted <LCH>
F:43-285/Domain: trypsin homology <TRY>
F:302-379/Region: proline-rich
F:22-210/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:25-154/Disulfide bonds: #status predicted
F:29-162/Disulfide bonds: #status predicted
F:73-89/Disulfide bonds: #status predicted
F:88-142,240/Active site: His, Asp, Ser #status predicted
F:177-246/Disulfide bonds: #status predicted
F:206-225/Disulfide bonds: #status predicted
F:236-266/Disulfide bonds: #status predicted

Query Match 20.5%; Score 481; DB 1; Length 421;
Best Local Similarity 41.0%; Pred. No. 4,5e-32;
Matches 100; Conservative 38; Mismatches 88; Indels 18; Gaps 5;

202 RVVGGEEASVDSWPMQVSIQY-----DKQHVGGSILDPHWLTAHCFRKHTDVENKVV 256
42 RIVGKAQHGAMPMWVSLQIFRYNSHYHCTGSSLNSRWLTAHCFVGKKNVDWRL 101

```

QY      257  RAGSJKL---SFSPELA-----VAKIIIEENPMYKPNIDALMLTQEPFSSGVPRP1 307
      : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db      102  VFGAEKIIYYGNKKPVKAPLQERHYVERKIIIEKYNATSEGNIDALVAETPIISGRFIRG 151

QY      308  CLPFFDEELTPAT-PLWIIIGWGFTTRKONGSKMSDILLQASVOYIDSTRCNADDAVQGEVTE 366
      : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db      162  CLPHFKAGLPKPSGSCWVAWGMYIEBKAPRPSIIIMEARVDLIDLCLNSTQWYNRVPQ 221

QY      367  KMKCMGIEPEGVDITQGDSDSGGFLMYQSDQ--WHVVGIVSKSGYCGCGSPSTPGVYTRVSAV 423
      : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db      222  TNVCAGYVVGKIDITQGDSDSGGFLMKDSKESAYVVVVGLTISWGVGALARRPGIYITATWPY 281

QY      424  INMI 427
      : : : : :
Db      282  INMI 285

```

Search completed: December 5, 2005, 13:53:51
Job time : 41 secs

Query Match	20.5%	Score 481	DB 1	Length 421
Best Local Similarity	41.0%	Pred. No. 4.5e-32		
Matches	100	Conservative	38	Mismatches 89
			Indels	18
			Gaps	5
QY	202	KVYGGEEASVDSPMVSYIQ----	DKQVCGGSLDPHWLVTAAHCRKTTDYPNNKY	256
DB	42	RIYGGAAAHGMAPMVVSYIQIRYNSHRHTCGGSLNSRWLTAAHCVGKNNVADWEL	101	

Db 169 PRPSVLLAEARVELLDLCSTQWYNGVSTWNCAGYPRGKIDTCGDSGGLMCRDN 248
Qy 393 -SDQWNVGVISMGVCGGSPSTPGYTKVSAIYLMWI 427
Db 249 ANSPFVVVGLTSMGVCARAKRPGIYTAWTWYLDWI 284

RESULT 13

KOHUP

plasma kallikrein (EC 3.4.21.34) precursor - human

N/Alternate names: kininogenin; plasma prekallikrein

C/Species: Homo sapiens (man)

C/Date: 13-Aug-1986 #sequence_revision 13-Aug-1986 #text_change 05-Oct-2004

C/Accession: A00921; A37939

R/Chung, D.W.; Fujikawa, K.; McMullen, B.A.; Davie, E.W.

Biochemistry 25, 2410-2417, 1986

A/Title: Human plasma prekallikrein, a zymogen to a serine protease that contains four

A/Reference number: A00921; MUID:86243359; PMID:3521732

A/Accession: A00921

A/Molecule type: mRNA

A/Residues: 1-638 <CHU>

A/Cross-references: UNIPROT:P03952; UNIPARC:UPI00000008AE; GB:M13143; NID:G190262; PIDN:

R/McMullen, B.A.; Fujikawa, K.; Davie, E.W.

Biochemistry 30, 2050-2056, 1991

A/Title: Location of the disulfide bonds in human plasma prekallikrein: the presence of

A/Reference number: A37939; MUID:91152016; PMID:1998666

A/Accession: A37939

A/Molecule type: protein

A/Residues: 20-27,40-46, 'X', 48, 'H', 50, 'X', 52-70, 'H', 75-76, 'X', 78-80, 103-113, 131-140, 141-

525,538-551,552, 'X', 564-567,573, 'X', 575-576,578-583, 'X', 585,592-604 <MCM>

A/Cross-references: UNIPARC:UPI00000172B4; UNIPARC:UPI00000172B4; UNIPARC:UPI00000172B4;

B49; UNIPARC:UPI00000172B4A; UNIPARC:UPI00000172B4B; UNIPARC:UPI00000172B4C; UNIPARC:UPI000

10000172B5; UNIPARC:UPI00000172B5; UNIPARC:UPI00000172B5; UNIPARC:UPI00000172B5; UNIPAR

C/Comment: This protein, synthesised in the liver, circulates as a noncovalent complex w

are linked by one or more disulfide bonds.

C/Comment: The zymogen is activated by factor XIIa, which cleaves the molecule into a li

C/Comment: The enzyme cleaves Lys-Arg and Arg-Ser bonds. It activates, in a reciprocal r

C/Comment: and may also play a role in the renin-angiotensin system by converting proteinin i

C/Genetics:

A/Gene: GDB:KLK3

A/Cross-references: GDB:127575; OMIM:229000

A/Map position: 4q35-4q35

C/Superfamily: plasma kallikrein/coagulation factor XI; trypsin homology

C/Keywords: blood coagulation; duplication; fibrinolysis; glycoprotein; hydrolase; infla

F/1-19/Domain: signal sequence #status predicted <Sig>

F/20-638/Product: plasma kallikrein #status predicted <MAT>

F/20-390/Domain: plasma kallikrein heavy chain #status predicted <HCH>

F/110-199/Domain: apple repeat <AP1>

F/200-289/Domain: apple repeat <AP2>

F/291-380/Domain: apple repeat <AP3>

F/391-621/Domain: plasma kallikrein light chain #status predicted <LCH>

F/391-621/Domain: trypsin homology <TRY>

F/21-104,47-57,111-194,137-166,141-147,201-284,227-256,231-237,292-375,322-328,383

F/127-308,396,453,494/binding site: carboxylate (Asn) (covalent) #status experimental

F/318-347,340-345/disulfide bonds: #status predicted

F/390-391/Cleavage site: Arg-Ile (coagulation factor XIIa) #status predicted

F/434,483,578/Active site: Arg, Asp, Ser #status predicted

Query Match 21.0%; Score 491.5; DB 1; Length 638;
Best Local Similarity 42.2%; Pred. No. 1e-32; Indels 16; Gaps 8;
Matches 100; Conservative 47; Mismatches 74

Qy 313 DEELTPATPLWITGCFYTKONGKMSDILLQASVOVDISTRGNADAYQG-EVTEKAMCA 371
Db 508 GDTSTIYTCWWTGWMFSSHEK-GEIONILQKNIPLVNTEEC--QKRYDYKILTRMCA 564
Qy 372 GPEGVDPICGDSGGLMYQ--SDQWNVGVISMGVCGGSPSTPGYTKVSAIYLMWI 427
Db 565 GYKEGKADCKDSSGGLVCKNMRLVGLITSMGECARRBQPGVYTKVAYMDWI 621

RESULT 14

JEO315

low-density lipoprotein receptor-related protein - mouse

C/Species: Mus musculus (house mouse)

C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004

C/Accession: JEO315

R/Tomita, Y.; Kim, D.; Magocsi, K.; Fujino, T.; Yamamoto, T.T.

J. Biochem. 124, 784-789, 1998

A/Title: A novel low-density lipoprotein receptor-related protein with type II membrane

A/Reference number: JEO315; MUID:98429596; PMID:9756624

A/Accession: JEO315

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-1113 <TOM>

A/Cross-references: UNIPROT:Q92319; UNIPARC:UPI0000029874; DDBJ:AB013874; NID:G3869144,

C/Superfamily: mouse low-density lipoprotein receptor-related protein; LDL receptor lig

F/337-372/Domain: LDL receptor ligand-binding repeat homology <LDL1>

F/374-408/Domain: LDL receptor ligand-binding repeat homology <LDL1>

F/410-445/Domain: LDL receptor ligand-binding repeat homology <LDL1>

F/447-482/Domain: LDL receptor ligand-binding repeat homology <LDL1>

F/468-682/Domain: LDL receptor ligand-binding repeat homology <LDL1>

F/684-720/Domain: LDL receptor ligand-binding repeat homology <LDL1>

F/723-757/Domain: LDL receptor ligand-binding repeat homology <LDL1>

F/869-1097/Domain: trypsin homology <TRY>

Query Match 20.9%; Score 488.5; DB 2; Length 1113;
Best Local Similarity 33.2%; Pred. No. 3.5e-32;
Matches 130; Conservative 52; Mismatches 148; Indels 61; Gaps 17;

Qy 69 IPRQOLQDELDCPLGDEEHCVKSPFEGPAVAVRLSKIRSTLOVDSATGNMFSACFDN 128

Db 736 VPRDLMCDGWVDCSDSDSEMGCVLTKNG-----NSSLLTVHKAKEH--HYCAGD 785

Qy 129 FTEALAEACRQMGYSKPTFAVEIGPPDDLVETENSQELRM---RNSG----- 179

Db 786 WRETLSQACKQKGLGEPSTKLI---PGQ-----EQQWLRLPYWMENLNGSTLOE 834

Qy 180 -----PCLSGSLVSLHCLA--CGK---SLKTPRVVGEBAVDSPWQVSTQYDKQ-HV 227

Db 835 LLYVRHSCPSRSEISLLCSKQDGRPPARMKRILLGKTSRGRPMQCSIQSEPSGHI 894

Qy 228 CGGSIIIDPHVVLTAACFRKHTDVFNWXYRAGSDKL--GSPESLAVAKIILIEFNPMY 284

Db 895 CGCVLAKKKWVLVAACFEGRBDAVAKVFGINNLDHSGFGQTRFVKTILL--HPRS 952

Qy 285 K---DNDIADLMKQPLFTSGVTRPICLPFPBELTPAPVITLIGVFKQNGKMSDIL 341

Db 953 RAYVDIDISVELSDINETSIVRPVCLPSPEYLEPDTYCYITG--HMGNKMPFL 1009

Qy 342 LQASVOVIDSTRGNADAY--QGEVTEKAMCAIGEGVDTQCGDSGGLMYQ--SDQW 397

Db 1010 QGEVEVRIIPLEQC---QSYFDMKTTINRMICAGYESGTVDSQMSGGLVCEVRGQWT 1066

Qy 398 VVGVISMGVCGGSPSTPGYTKVSAIYLMWI 427

Db 1067 LRELITSMGVCVPSKVLGPGVSNVSPYVGM 1097

RESULT 15

S11674

acrosin (EC 3.4.21.10) precursor - human

C/Species: Homo sapiens (man)

C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004

C/Accession: S11674; S23499; S12063; A61022; S03330

A:Residues: 1-418 <KLE>
A:Cross-references: UNIPARC:UPI000016CBCE, GB:X52466; NID:g94987; PIDN:CAA36704.1; PID:
C/Superfamily: acrosin; trypsin homology
C:Keywords: glycoprotein; hydrolase; serine proteinase
F./39-/Domain: trypsin homology <TRY>
F./18-/Binding site: carbohydrate (Asn) (covalent) #status predicted
F./21-/Disulfide bonds: #status predicted
F./25-/Disulfide bonds: #status predicted
F./70-/Disulfide bonds: #status predicted
F./85,/139,/238/Active site: His, Asp, Ser #status predicted
F./175-/244/Disulfide bonds: #status predicted
F./207-/223/Disulfide bonds: #status predicted
F./234-/264/Disulfide bonds: #status predicted

Query Match 21.5%; Score 503; DB 2; Length 418;
Best Local Similarity 39.1% Pred. No. 6.7e-34;
Matches 108; Conservative 36; Mismatches 90; Indels 42; Gaps 7;

OY 179 GPLSSLSVLHCLACKSLKT-----PRVGGEEASVDSPMPQVSIOY-----DIQ 225
DB 23 GP-----GLTRONSGAGTIVISGSAHVGMPPWWSLQIFTSNSRR 67

OY 226 HVCGSILDPHWYLTAACRKHDFENMK-----VRAGSDXGSFP--SLAVAKII 276
DB 68 HACSGLINSHWULTAHCPDNKKVYDMRLVFQAQEIEYGNNKYKEOEERYOKIYI 127

OY 277 IEFNPYPKDNDIALMKLQPPLTFSGTVRPICLPFFDELTAPATP-LWIIGMTFKONG 334
DB 128 HEKVNVVTBENDIALALNTPTPVTCGNFIGPCCLPFHKAGPRPKIPTCTCVTMGYIKREA 187

OY 335 GKMSDILLQASVOYIDSTRCADDAAYOGVEYTEKMKCAGIPESGVDTCCGDSSGPLMYQS 394
DB 188 PRPSPILMEARVLDIDLDCNSTOMYNGRVSTNCAGPEBKIDTPCCDSGPGIMCRDN 247

OY 395 ---QHWHVGIVSMVGCYGGCSPTPGVYTCKVASLIAMI 427
DB 248 ARQPFVVVGITSWGVCARAKRPGVIYAFTWDILMI 283

RESULT 11
S18407
acrosin (EC 3.4.21.10) precursor - rat
N/Ccontains: proacrosin
C/species: Rattus norvegicus (Norway rat)
C/date: 31-Dec-1993 #sequence revision 31-Dec-1993 #text_change 09-Jul-2004
C/accession: S18407; S30037; A56620
R/Klennm, U.; Flake, A.; Engel W.
Biochim. Biophys. Acta 1090, 270-272, 1991
Afiliate: Rat sperm acrosin: cDNA sequence, derived primary structure and phylogenetic or
A/reference number: S18407; MUID:92031708; PMID:1932123
A/accesion: S18407
A/molecule type: mRNA
A/residues: 1-437 <KLE>
A/cross-references: UNIPROT:P29293; UNIPARC:UPI00001253A0; EMBL:X59254
submitted to the EMBL Data Library, April 1991
A/reference number: S30037
A/accesion: S30037
A/molecule type: mRNA
A/residues: 1-254,'LCDR','259','DHET','264','GRLC','269-437'<KLE2>
A/cross-references: UNIPARC:UFI0000170BZ3; EMBL:X59254; NID:g957282; PIDD:CAA1947.1; PIDD:
R.Klemmling, H.; Flake, A.; Adham, I.M.; Radtke, J.; Engel, W.
DNA Seq. 2, 57-60, 1991
A/filte: Exon-intron structure and nucleotide sequence of the rat proacrosin gene.
A/reference number: A56620; MUID:92199245; PMID:1802037
A/accesion: A56620
A/status: preliminary
A/molecule type: DNA; mRNA
A/residues: 1-254,'LCDR','259','DHET','264','GRLC','269-437'<KRB>
A/cross-references: UNIPARC:UPI0000170B23
A>Note: sequence modified after extraction from NCBI backbone
C/Superfamily: acrosin; trypsin homology

C;Keywords: glycoprotein; hydrolase; serine proteinase; sperm; zymogen

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-43/Domain: acrosin light chain #status predicted <LCH>

F:43-286/Domain: trypsin homology <TRY>

F:44-437/Domain: acrosin heavy chain #status predicted <HCH>

F:22_21/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:152-155,29-163/disulfide bonds: #status predicted

F:74-90/disulfide bonds: #status predicted

F:189,143,241/Active site: His, Asp, Ser #status predicted

F:178-247/dissulfide bonds: #status predicted

F:210-226/disulfide bonds: #status predicted

F:237-267/disulfide bonds: #status predicted

```
Query Match      21.4%; Score 501.5; DB 2; Length 437;  
Best Local Similarity   42.7%; Pred. No. 9,4e+34;  
Matches    105; Conservative     35; Mismatches    85; Indels     21; Gaps     6;  
  
QY       202 RVVGGEASVDSWPMOWSIQY-----DKOHVCGSLIDPHHWTLAAHCERKHTDVFMWK 255  
         |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:  
DB        42 RIVGGOTSRMAMPWMVSLOIFTSNRSRRTHACGGLNSHWLTAAHCDFNKKKRYDWR 101  
         :::::::::::::::::::::  
QY       256 -----VRAGSDKLGSFPSLA--VAKIIEFNMYPKNDIALMKQLPFLTSGTVRP 306  
         |  
DB       102 LVFGAHEIETYGNNKPVEKEPDGERVVOKLVHEKNATLEGNDIALKTPTVTCDPFGP 161  
         :|:::||::|  
QY       307 ICLPPFDIELTPATP--LMIIWGFTFKONGCKMSDLLQSVOVIYSTRCNADAYOGEV 364  
         |||::|||::|||::|||::|||::|||::|||::|||::|||::|||:  
DB       162 GLPLPHF-KSGCPRIPTHCTCVTTGWGYIKDNA PRSPVLMEARVDIDLDCNSTGMWNNGV 220  
         ::::|  
QY       365 TEKMNCAGIPEGGVDPICQGDSGGPELWYSOQ---MHNVGIYSWGVGCCGSPITPVYTAKS 421  
         ::::|  
DB       221 TSTNYCAGAPBEGRKITPCQGDSSGPLMCRTDRROPFIIVGITSWGVGCARAKRPGVYATW 280  
         ::|||  
QY       422 AYLNMI 427  
         |||:  
DB       281 DYLDWI 286  
         ::|||  
  
RESULT 12  
S29599  
acrossin (EC 3.4.21.10) precursor - guinea pig (fragment)  
C/Species: Cavia porcellus (guinea pig)  
C/Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004  
C/Accession: S29599  
R/Gerton, G.L.; Hoff, H.B., Baba, T.  
submitted to the EMBL Data Library, May 1992  
A/Description: The amino acid sequence of guinea pig proacrosin deduced from its cDNA se-  
quence.  
A/Reference number: S29599  
A/Accession: S29599  
A/Molecule type: mRNA  
A/Residues: 1-421 <GER>  
A/Cross-references: UNIPROT:O60491; UNIPARC:UPI00000E5c67; EMBL:Z12153; NID:g49559; PIDN  
B:Superfamily: acrossin; trypsin homology  
C/Keywords: glycoprotein; hydrolase; serine proteinase  
F:/41-284/Domain: trypsin homology <TRY>  
  
Query Match      21.1%; Score 493.5; DB 2; Length 421;  
Best Local Similarity   38.0%; Pred. No. 4,1e+33;  
Matches    105; Conservative     44; Mismatches    84; Indels     43; Gaps     8;  
  
QY       179 GFC-----LSGLSVSLHLCLACGSKLTPRVVGEEASVDSWPMOWSIQY-----DKQ 225  
         |||:  
DB       25 GFCGFPRONLGSAV-----RLIGGTQAQPAGAWMVWSLOIFAHHANNRY 69  
         ::|||::|||::|||::|||::|||::|||::|||::|||::|||:  
QY       226 HYCCGSIILPHWTLLTAHCERKHDTVENMKVRAGSDKL----GSPSLA-----VAKIII 276  
         |||::|||::|||::|||::|||::|||::|||::|||::|||:  
DB       70 HACGGILLNSHWULTAHCFDSKKKYDMRLFCAEBELTYGNKPYRAAPLOBERYIVEKIYA 129  
         ::|||::|||::|||::|||::|||::|||::|||::|||:  
QY       277 IEFNPNYPKDNIALMLKQLPFLTFSGTVRPICLPFPDEILTPTAP--LMIIWGFTKONG 334  
         |||::|||::|||::|||::|||::|||::|||::|||:  
DB       130 HEKYNINVNENGINDLALKITPIPVSCGPFIGRGCLPTF-RAGSPRKIQTCYVAVMGVIREKA 188  
         ::|||::|||::|||::|||::|||::|||::|||:  
QY       335 GKMSDILLQASVOVIDSTRCNADDAVAGEVTEKMMCAIGPEGCVDTCCGDSGSLMTQ-- 392
```

F:21-104,47-77,51-57,111-194,137-166,141-147,201-284,227-256,231-237,292-375,318-347,322
F:127,215,308,396,434/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:434,483,578/Active site: His, Asp, Ser #status predicted

Query Match

21.7%; Score 508; DB 1; Length 638;

Best Local Similarity 41.0%; Pred. No. 4,3e-34;

Matches 112; Conservative 45; Mismatches 82; Indels 34; Gaps 10;

QY 174 MRNSGPGCLSGSLVSHCLACGKSLKTP-----RVNGGEASVDSMPWQVSIQ---Y 222

DB 364 MQSSSG-----YSRLCKLVDSPTCTTKINARIVGVTNASTGEMWQVSLQVKLV 413

QY 223 DKQHVCGGSLTDHFWLTLAHCFR--KHTDVFMMKTRAG---SDKLGSFPSLAIAKII 276

DB 414 SQTHLCGGSLIGRWLTLAHCFDGIYDPV--WRIYGGILSLSEITKETPSSRIKELII 471

QY 277 IEFNPMYPRKNDIALMKLOPLTFSGTVRPICLPFDEELTPATPLMIIGMGTKONGK 336

DB 472 HQRYKXSEGVYDIALIKLOPLNTYEFQFICLPKRDNTITNTNCVWGMGTKEQ-GR 530

QY 337 MSDILLQASVQVDSITRCNADADAYQGEVTER--NMCAIGPEGVDTCCGDSGGLMYQ-SD 394

DB 531 TQNILOKATRIPLVPENEC--QKRYRDYVINKQMICAGYKRGCTDACKGDSGGLVCKHSG 588

QY 395 QMHWGVIVSWGYCCGSPSTPGVYTKYSATLAWI 427

DB 589 RMQLVGITSMWEGCGRRKDPGVYTKVSEYWDWI 621

RESULT 8

JX0172

acrosin (EC 3.4.21.10) precursor form 3 - mouse

C/Species: Mus musculus (house mouse)

C/Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004

C/Accession: JX0172; JX0138

R:Watanabe, K.; Baba, T.; Kashiwabara, S.; Okamoto, A.; Arai, Y.

J. Biochem. 109, 828-833, 1991

A>Title: Structure and organization of the mouse acrosin gene.

A/Reference number: JX0172; MUID:92041732; PMID:1939002

A/Accession: JX0172

A/Molecule type: DNA

A/Residues: 1-436 <WAT>

A/Cross-references: UNIPROT:P23578; UNIPARC:UPI000021F79; GB:S66245; NID:9238706; PIDN:R:Kashiwabara, S.; Baba, T.; Takada, M.; Watanabe, K.; Yano, Y.; Arai, Y.

J. Biochem. 108, 785-791, 1990

A>Title: Primary structure of mouse proacrosin deduced from the cDNA sequence and its gene

A/Reference number: JX0138; MUID:91185335; PMID:2127931

A/Accession: JX0138

A/Molecule type: mRNA

A/Residues: 4-436 <KAS>

A/Cross-references: UNIPARC:UPI000016CBD; GB:D00754; NID:9220322; PIDN:BAA00651.1; PID: C/Comment: Acrosin is an acrosomal protease that plays an important role in the initial e-binding activity.

C/Genetics: 26/2; 95/2; 190/1; 238/3

A/Intron: 26/2; 95/2; 190/1; 238/3

C/Superfamily: acrosin; trypsin homology

C/Keywords: glycoprotein; hydrolase; serine proteinase; sperm

F:1-19/Domains: signal sequence #status predicted <SIG>

F:20-42/Product: acrosin light chain #status predicted <ALC>

F:43-321/Product: acrosin heavy chain #status predicted <AHC>

F:43-286/Domains: trypsin homology <TRY>

F:22.211/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:25-155,29-162,74-90,178-226,237-267/Dissulfide bonds: #status predicted

F:89,143,241/Active site: His, Asp, Ser #status predicted

F:235/Binding site: substrate (Asp) #status predicted

Query Match

21.6%; Score 505; DB 2; Length 436;

Best Local Similarity 40.9%; Pred. No. 4,8e-34;

Matches 106; Conservative 37; Mismatches 90; Indels 26; Gaps 7;

QY 194 CG-----KSLKTPRVYVGEASVDSMPWQVSIQ-----DKQHVCGGSLTDHFWLTLA 242

DB 29 CGLRFRNSQAGTRIVSGSADLGAPMNVSLQIFTSNRSRYHACGSLNSHWLTLA 88

QY 243 HCFRKHITDVFNK-----VRAGSDKLGSFPSLA--YAKIIIEFNPMYPRKNDIALMK 293

DB 89 HCFDNKKKYDWRVLVGAQIEYGRNKPVEKQGRYVQKIVIHKKYNNVTEGNDIALK 148

QY 294 LQPLTFSTVAPRICLPFDEELTPATP--LMIIGMGTKONGKMSDILLQASVQVDS 351

DB 149 ITPPVTCGNFIPGCPCLPHF--KAGPQIPIHTCVVTGMYIKERAPRPSPLVMEARVLDL 207

QY 352 TRCNADADAYQGEVTERKMGAGIPEGVDTCCGDSGGLMYQSD---QMHWGVIVSWGYCC 408

DB 208 DLCSNTQWTVNGVSTTNVCAGYPEBKIDTCQDSGGLMCRDNDVSPFVVGITSMWGYC 267

QY 409 GGPSTPGVYTKYSATLAWI 427

DB 268 ARAKRPGVYATATWLDWI 286

RESULT 9

T30337

polyprotein - African clawed frog

C/Species: Xenopus laevis (African clawed frog)

C/Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004

C/Accession: T30337

R:Yang, J.C.; Lindsey, L.L.; Hedrick, J.L.

submitted to the EMBL Data Library, March 1998

A/Description: cDNA cloning of ovocytinase, a chymotrypsin-like protease released from X

A/Reference number: Z20829

A/Accession: T30337

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-1524 <YAN>

A/Cross-references: UNIPROT:Q91674; UNIPARC:UPI00000FBA76; EMBL:U81290; NID:92581640; P

C/Superfamily: trypsin related polyprotein; trypsin homology

Query Match

21.5%; Score 504.5; DB 2; Length 1524;

Best Local Similarity 44.0%; Pred. No. 2,4e-33;

Matches 109; Conservative 38; Mismatches 86; Indels 15; Gaps 8;

QY 194 CGKSLKT-----PRVNGGEASVDSMPWQVSIQDKQHVCGGSLTDHFWLTLAHCFRKH 248

DB 570 CGMAWTRKMWLPRVNGGEASVDSMPWQVSIQDKQHVCGGSLTDHFWLTLAHCFRKH 629

QY 249 TDVFNKTRAGS--DKLGSFPSLAIAKIIIEFNPMYPRK---DNIDIALMKLOPLTFSGTV 304

DB 630 EBSY-WYIAGDHRMNLNESTEQIRNITIRIHNYNSETYDNDIALYLSEPLDINPV 688

QY 305 RPTCLPFDEELTPATPLMIIGMGTKONGKMSDILLQASVQVDSITRCNADADAYQGEV 364

DB 689 RPYCLPEBEVLTPTASVCVTVGMGTADGQPALG-LQQLPLTIDSLICNT--SYSGEL 746

QY 365 TEKMKCAIIPBG--GVDTCCGDSGGLMYQSD---DMHWGVIVSWGYCCGSPSTPGVYTKYS 421

DB 747 TDHMLCAGFPSSKEXKDCQDSGGLVYQNEKEQFSIVGLVSWGCGCRVSPGVYTKYR 806

QY 422 AYLMWTVN 429

DB 807 LFTWYIQN 814

RESULT 10

A37344

acrosin (EC 3.4.21.10) precursor form 1 - mouse

C/Species: Mus musculus (house mouse)

C/Date: 19-Mar-1993 #sequence_revision 19-Mar-1993 #text_change 22-Jun-1999

C/Accession: A37344

R:Klemm, U.; Meier, W.M.; Tsounisidou, S.; Adam, I.M.; Willison, K.; Engel, W.

A>Title: Mouse proacrosin: cDNA sequence, primary structure and postmeiotic expression

A/Reference number: A37344; MUID:90255839; PMID:2111255

A/Accession: A37344

A/Status: preliminary

A/Molecule type: mRNA

Qy	190	HC--LACCGL-----KT	PRVVGEEASVDMPMVSIQYDKRHYCGGSLILDPHWLTAAH	243
Db	781	QCNHNSCGKQVAQAEVSPIKVGGNDSSREGAMPWVALTYNQQLLCGASLVSRDMLVSAAH		840
Qy	244	C-FRKHTDVFNMKVRAGSDKLGSF--PSLA	VAKIIIEFNMY--PKNDI	ALMKTLOP 297
Db	841	CYYGNLEBSKKWKAILGLHMTSNLTSPOIVTRLDEIVINPHYNRRRDSIDIANHTEERK		900
Qy	298	LTFSGTVRPICLPFEDBELTPATPLMWITIGMGFTKQNGSKMSDILQAASVOYIDSTRCANAD		357
Db	901	VNYTDVIQPCLPFEENQVFPFGRCISLAGMSKVLYQGSP-ADILQEAADVPLLSNRKCQ-Q		958
Qy	358	DAYOGEVTERKMNCAGIPBEGVDTCQGDGGGLM-YQSDQMWHVGI	VMSSGYGGGSPSTPGV	416
Db	959	QMPENVNITENNMCAGEYGIDSQGDGSGPLMCLLENRRILLAGVTISFGYQCALENRPGV		1018
Qy	417	YTKVASAYLNMVI	427	
Db	1019	VARVPKFTEMI	1029	

RESULT 6

plasma kallikrein [EC 3.4.21.34] Precursor- rat
 N:Alternate names: Fletcher factor; kininogenin; serum kallikrein
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 05-Oct-2004
 C:Accession: A39180; A33320; S06851; I53041; S06852
 R:Seidah, N.G.; Ladenheim, R.; Mbikay, M.; Chretien, M.; Seidah, N.G.
 Biochemistry 30, 1658-1635, 1991
 A:Title: Gene structure and chromosomal localization of plasma kallikrein.
 A:Reference number: A39180; MUID:91129236; PMID:1993180
 A:Accession: A39180
 A:Molecule type: DNA
 A:Residues: 1-638 <BEA>
 A:Cross-references: UNIPROT:P14272; UNIPARC:UPI000004ECC9; GB:J05315
 A>Note: the authors translated the codon GAG for residue 81 as Gln
 R:Seidah, N.G.; Ladenheim, R.; Mbikay, M.; Hamelin, J.; Lutfalla, G.; Rougeon, F.; Lazure
 DNA 8, 563-574, 1989
 A:Title: The CDNA structure of rat plasma kallikrein.
 A:Reference number: A33320; MUID:90091743; PMID:2598771
 A:Accession: A33320
 A>Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-638 <SEI>
 A:Cross-references: UNIPARC:UPI000004ECC9; GB:M30282; NID:9205010; PIDN:AAA4463.1; PID:
 A>Note: part of this sequence, including the amino ends of both the heavy and light chain
 R:Paquin, J.; Benjamel, S.; Sawyer, N.; Lazure, C.; Chretien, M.; Seidah, N.G.
 Biochim. Biophys. Acta 999, 103-110, 1989
 A:Title: Rat plasma kallikrein: purification, NH(2)-terminal sequencing and development
 A:Reference number: S06851; MUID:90089457; PMID:2597701
 A:Accession: S06851
 A:Molecule type: protein
 A:Residues: 20-45;391-413 <PAQ>
 A:Cross-references: UNIPARC:UPI0000172B57; UNIPARC:UPI0000172B58
 R:Seidah, N.G.; Ladenheim, R.; Mbikay, M.; Hamelin, J.; Lutfalla, G.; Rougeon, R.; Lazure
 DNA Cell Biol. 8, 563-574, 1989
 A:Title: The CDNA structure of rat plasma kallikrein.
 A:Reference number: I53041
 A:Accession: I53041
 A>Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-638 <RES>
 A:Cross-references: UNIPARC:UPI000004ECC9; GB:M58590; NID:9206721; PIDN:AAA42069.1; PID:
 C:Comment: This protein, synthesized in the liver, circulates as a noncovalent complex with
 are linked by one or more disulfide bonds.
 C:Genetics:
 A:Gene: PK
 C:Superfamily: plasma kallikrein/coagulation factor XI; trypsin homology
 C:Keywords: Blood coagulation; duplication; fibrinolysis; glycoprotein; hydrolase; infla
 F:1-19/Domain: signal sequence #status predicted <Sig-

F:20-390/Product: plasma kallikrein heavy chain #status experimental <MART>
F:20-109/Domain: apple repeat <AP1>
F:110-199/Domain: apple repeat <AP2>
F:200-289/Domain: apple repeat <AP3>
F:1291-380/Domain: apple repeat <AP4>
F:391-638/Product: plasma kallikrein light chain #status experimental <MART>
F:391-631/Domain: trypsin homology <TRY>
F:21-104,47-77,51-57,111-194,137-166,141-147,201-284,227-256,231-237,292-375,318-347,322-343,345-350,351-355,356-360,361-365,366-370,371-375,376-380,381-385,386-390,391-395,396-400,401-405,406-410,411-415,416-420,421-425,426-430,431-435,436-440,441-445,446-450,451-455,456-460,461-465,466-470,471-475,476-480,481-485,486-490,491-495,496-500,501-505,506-510,511-515,516-520,521-525,526-530,531-535,536-540,541-545,546-550,551-555,556-560,561-565,566-570,571-575,576-580,581-585,586-590,591-595,596-600,601-605,606-610,611-615,616-620,621-625,626-630,631-635,636-640,641-645,646-650,651-655,656-660,661-665,666-670,671-675,676-680,681-685,686-690,691-695,696-700,701-705,706-710,711-715,716-720,721-725,726-730,731-735,736-740,741-745,746-750,751-755,756-760,761-765,766-770,771-775,776-780,781-785,786-790,791-795,796-800,801-805,806-810,811-815,816-820,821-825,826-830,831-835,836-840,841-845,846-850,851-855,856-860,861-865,866-870,871-875,876-880,881-885,886-890,891-895,896-900,901-905,906-910,911-915,916-920,921-925,926-930,931-935,936-940,941-945,946-950,951-955,956-960,961-965,966-970,971-975,976-980,981-985,986-990,991-995,996-1000,1001-1005,1006-1010,1011-1015,1016-1020,1021-1025,1026-1030,1031-1035,1036-1040,1041-1045,1046-1050,1051-1055,1056-1060,1061-1065,1066-1070,1071-1075,1076-1080,1081-1085,1086-1090,1091-1095,1096-1100,1101-1105,1106-1110,1111-1115,1116-1120,1121-1125,1126-1130,1131-1135,1136-1140,1141-1145,1146-1150,1151-1155,1156-1160,1161-1165,1166-1170,1171-1175,1176-1180,1181-1185,1186-1190,1191-1195,1196-1200,1201-1205,1206-1210,1211-1215,1216-1220,1221-1225,1226-1230,1231-1235,1236-1240,1241-1245,1246-1250,1251-1255,1256-1260,1261-1265,1266-1270,1271-1275,1276-1280,1281-1285,1286-1290,1291-1295,1296-1300,1301-1305,1306-1310,1311-1315,1316-1320,1321-1325,1326-1330,1331-1335,1336-1340,1341-1345,1346-1350,1351-1355,1356-1360,1361-1365,1366-1370,1371-1375,1376-1380,1381-1385,1386-1390,1391-1395,1396-1400,1401-1405,1406-1410,1411-1415,1416-1420,1421-1425,1426-1430,1431-1435,1436-1440,1441-1445,1446-1450,1451-1455,1456-1460,1461-1465,1466-1470,1471-1475,1476-1480,1481-1485,1486-1490,1491-1495,1496-1500,1501-1505,1506-1510,1511-1515,1516-1520,1521-1525,1526-1530,1531-1535,1536-1540,1541-1545,1546-1550,1551-1555,1556-1560,1561-1565,1566-1570,1571-1575,1576-1580,1581-1585,1586-1590,1591-1595,1596-1600,1601-1605,1606-1610,1611-1615,1616-1620,1621-1625,1626-1630,1631-1635,1636-1640,1641-1645,1646-1650,1651-1655,1656-1660,1661-1665,1666-1670,1671-1675,1676-1680,1681-1685,1686-1690,1691-1695,1696-1700,1701-1705,1706-1710,1711-1715,1716-1720,1721-1725,1726-1730,1731-1735,1736-1740,1741-1745,1746-1750,1751-1755,1756-1760,1761-1765,1766-1770,1771-1775,1776-1780,1781-1785,1786-1790,1791-1795,1796-1800,1801-1805,1806-1810,1811-1815,1816-1820,1821-1825,1826-1830,1831-1835,1836-1840,1841-1845,1846-1850,1851-1855,1856-1860,1861-1865,1866-1870,1871-1875,1876-1880,1881-1885,1886-1890,1891-1895,1896-1900,1901-1905,1906-1910,1911-1915,1916-1920,1921-1925,1926-1930,1931-1935,1936-1940,1941-1945,1946-1950,1951-1955,1956-1960,1961-1965,1966-1970,1971-1975,1976-1980,1981-1985,1986-1990,1991-1995,1996-2000,2001-2005,2006-2010,2011-2015,2016-2020,2021-2025,2026-2030,2031-2035,2036-2040,2041-2045,2046-2050,2051-2055,2056-2060,2061-2065,2066-2070,2071-2075,2076-2080,2081-2085,2086-2090,2091-2095,2096-2100,2101-2105,2106-2110,2111-2115,2116-2120,2121-2125,2126-2130,2131-2135,2136-2140,2141-2145,2146-2150,2151-2155,2156-2160,2161-2165,2166-2170,2171-2175,2176-2180,2181-2185,2186-2190,2191-2195,2196-2200,2201-2205,2206-2210,2211-2215,2216-2220,2221-2225,2226-2230,2231-2235,2236-2240,2241-2245,2246-2250,2251-2255,2256-2260,2261-2265,2266-2270,2271-2275,2276-2280,2281-2285,2286-2290,2291-2295,2296-2300,2301-2305,2306-2310,2311-2315,2316-2320,2321-2325,2326-2330,2331-2335,2336-2340,2341-2345,2346-2350,2351-2355,2356-2360,2361-2365,2366-2370,2371-2375,2376-2380,2381-2385,2386-2390

Query Match	21.8%;	Score 511;	DB 1;	Length 638;
Best Local Similarity	32.0%;	Pred. No. 2.4e-34;		
Matches 128;	Conservative 62;	Mismatches 106;	Indels 104;	Gaps 17;

```

0Y      ::::-----ELDCPL-GED--EEHCVKSF-----PEGP 98
64 QPILFIPRKQLCDG-----

```

00 АВАИМІ СУВРОВА ОУТ ДОАТОЧНЕСОСЕРДНЕТЕАЛ АЕТА СРОМСЫСКРЕПЕРАВЕІСДО 158

Db 346 KCSLRSTDGSPRITYEAGS-----SGYS----- 371

QY 159 DLDVEITENSQELRMNSSGPECLSGSLVSLHCLACGSKLTPRVVGEEASVDSWPQV 218

```

Db      3/2-LRCKKVESSD-----CINLNA-KIVGINSUGEMFY 300

```

407 SLOVKLVSONHMGSGSIIGROWILTAHCFDGIPIYDV--WRIYGLINLSEITNKTPFS 464

270 AVAKIIIEFNPMYKNDIALMKIQPLTSGTVRPICLPFDEELTPATPLMIIGWG 329

Db 465 SIKELIHOKYKMSGYDALIKLOTPLNTEFQKPICLPSKADNTIYINCWGTGNGY 524

```

581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000

```

QY 389 LMYQ-SDQMHVGIIVSMGYCGGSPITPGVYTKVSAYLNIWI 427

Db 582 LVCKHSGRQGLVGITSMGEGCARKEQPGVYTKVAEYIDWI 621

RESULT 7

plasma kallikrein (EC 3.4.21.34) precursor - mouse
C. Species: Mus musculus (house mouse)

C/Accession: A36557	#sequence_revision 30-Sep-1992	#rev_change 03 Oct 2003
---------------------	--------------------------------	-------------------------

DNA Cell Biol. 9, 737-748, 1990

A/Accession: A36557

A;Residues: 1-638 <SEI>
A;Cross-references: UNIPROT:P26262; UNIPARC:UPI00000277BC; GB:M58588; NID:g200

C;Comment: This protein, synthesized in the liver, circulates as a noncovalent C₂ dimer. The zymogen is activated by factor XIIa, which cleaves the molecule.

C: Superfamily: plasma kallikrein/coagulation factor XI; trypsin homology

F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-390/Product: plasma kallikrein heavy chain #status experimental <HCH>

```
F:\10-199\Domain: apple repeat <AP2>  
F:\20-107\Domain: apple repeat <AP2>  
F:\300-300\Domain: apple repeat <AP2>
```

F,291-380/Domain: apple repeat <AP4>
F,391-638/Product: plasma kallikrein light chain #status experimental <LCH>

F:801-1030/Domain: trypsin homology <TRY>
 F:116,147,170,194,233,263,264,404,456,466,519,550,646,698,722,741,762,864,903,965/Bindin
 F:788-912,826-842,926-973,957-972,983-1011/Disulfide bonds: #status predicted
 F:841,892,987/Active site: His, Asp, Ser #status predicted

Query Match 24.4%; Score 571; DB 1; Length 1035;
 Best Local Similarity 35.0%; Pred. No. 4,6e-39;

Matches 133; Conservative 67; Mismatches 138; Indels 42; Gaps 15;

```

QY 69 IPRKOLCDDELDCPLGDEDEHCVKSPPEGPAAVRLSKDRSTLQVLDATGWMFSAFCPN 128
DB 672 IFLVNLCDGFPHCKGSDDAHCVRLE-NGTTSSGLVQFR--IQSI-----MHVACAEV 722
QY 129 FTEALAEATACRQMGY----SSKPTFAVEIGPDODLVE---ITENSQELMRNMSGP 180
DB 723 WTTQSLSDVQCQLLGLGTGNSVPTF-STGGPFVNLNTAPNSGLITPESQ----- 772
QY 181 CLSGSLVSLHLC--LACGSKLKT---PRVYGEBEASVDSMPQVSIQYKQHYCGGSLID 234
DB 773 CLEDSLILLLQCNVKSCKKLVTQEVSPKIVGSDSREGAMPVVALYPDDQCGASIVS 832
QY 235 FHWVLTAAHC-PRKHIDVENMKVRAGSDKLGSF--PSLAVAKIIEFPMY---PKMD 288
DB 833 RMVLSAAHCVCYGRNMEPSKMAVGLHMASNLTSFQITRLIDQVINPHNKRKKN 892
QY 289 IALMKLQPLPFGSVTRPICLPFDEELTPATPLMIIGWFTKQNGKMSDILLQASVQV 348
DB 893 IAMHMLEMKVNTDYIQPLCLPEBNQVFPGRICSLAGGALILYQ-GSTADVLQEDAVPL 951
QY 349 IDSTCNADDAVQGEVTEKMCAGIPREGVDTQCGSDGSLMYQ-SDQMHVVGIVSWG 407
DB 952 ISNECKQ-QQMPPEYNTENMVCAGYERAGVDSQCGDSGGLMCOENRWLLAGVTSFGY 1010
QY 408 CGGSPGPGVYTKVSAVLMWI 427
DB 1011 CALPNRPGYARVPRTEWI 1030

```

RESULT 4

S33777
 hepsin (EC 3.4.21.-) - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 05-Oct-2004
 C:Accession: S33777; S32013
 R:Farley, D.; Raymond, F.; Nick, H.
 Biochim. Biophys. Acta 1173, 350-352, 1993
 A:Title: Cloning and sequence analysis of rat hepsin, a cell surface serine proteinase.
 A:Reference number: S33777; MUID:93305733; PMID:8318546
 A:Accession: S33777
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-416 <FAR>
 A:Cross-references: UNIPROT:Q05511; UNIPARC:UPI00004ECDC9; EMBL:X70900; NID:957928; PIDN
 C:Keywords: hydrolase; liver; serine proteinase; transmembrane protein
 F:12-44/Domain: transmembrane #status predicted <TM>
 F:162-399/Domain: trypsin homology <TRY>
 F:187-203,280-358,321-337,348-380/Disulfide bonds: #status predicted
 F:202,256,352/Active site: His, Asp, Ser #status predicted

Query Match 24.3%; Score 568.5; DB 1; Length 416;
 Best Local Similarity 30.5%; Pred. No. 2.5e-39;
 Matches 137; Conservative 64; Mismatches 155; Indels 93; Gaps 11;

```

QY 19 KPRIMEPRFKGPIITALLSLASIIYVVLIKYILDKRYFLCGQPLHFRPKOLCDGE 78
DB 14 RPKVAALT---VGTLLFLFGIGASMAIYITILLR----- 44
QY 79 LDCPLGDEDEHCVKSPPEGPAAVRLSKDRSTLQVLDATGWMFSAFCPNFTEALATAC 138
DB 45 -----SDQ-----PLYQVQLSPGDSRLVLVDKTEGTWRLLCSSRSNARVAGIGC 89
QY 139 RQMGY-----SSKPTFAVEIG---PDODLVEITENSQELMRNMSG 178
DB 139 RQMGY-----SSKPTFAVEIG---PDODLVEITENSQELMRNMSG 178

```

```

DB 90 EEMGFLRALANSELDVRTAGANGTSGFPCVEGGLPLAQRLLDVISVCD----- 138
QY 179 GPCLSGSLVSLHCLACG-KSLKTPVVGEBEASVDSMPQVSIQYDKQHYCGGSLIDPMH 237
DB 139 --CPRGRLTATQCQCGCRKLPVDRIVGGQSSLSKRWMPVSLRYDGHLLCGSLISGDM 196
QY 238 VLTAAHCERKATDVN-WKVRAGSDKLGSFSLAVAKIIEFPMY-----KOND 288
DB 197 VLTAAHCERKATDVN-WKVRAGSDKLGSFSLAVAKIIEFPMY-----KOND 288
QY 289 IALMKLQPLPFGSVTRPICLPFDEELTPATPLMIIGWFTKQNGKMSDILLQASVQV 348
DB 257 IALVHLSSTLPLETEYIQVCLPAAQALVDGKVCVTWTGWT-OPYGQAVVLDQEARVPI 315
QY 349 IDSTCNADDAVQGEVTEKMCAGIPREGVDTQCGSDGSLMYQ-----SDQMHVVGIVS 403
DB 316 ISNEVNSPDRFGNOIKTKMFCAGIPREGIDIDCQDSDSGHFCEDRISGTSRMRCLGIVS 375
QY 404 WGYGCGSPGPGVYTKVSAVLMWIYVWK 432
DB 376 WGTGALARKPGVYTKVLDIFREWIFQAIK 404

```

RESULT 5

A53663
 enteropeptidase (EC 3.4.21.9) precursor [validated] - pig
 N:Alternate names: enterokinase
 C:Species: Sus scrofa domestica (domestic pig)
 C:Date: 07-Oct-1994 #sequence_revision 09-Aug-1996 #text_change 28-Apr-2003
 C:Accession: A53663
 R:Matsumura, M.; Ichinose, M.; Yahagi, N.; Kakel, N.; Tsukada, S.; Miki, K.; Kurokawa,
 J. Biol. Chem. 269, 19976-19982, 1994
 A:Title: Structural characterization of porcine enteropeptidase.
 A:Reference number: A53663; MUID:94327548; PMID:8051081
 A:Accession: A53663
 A:Molecule type: mRNA
 A:Residues: 1-1034 <MAT>
 A:Cross-references: UNIPARC:UPI0000172B0C; GB:D30799; NID:9505122; PIDN:BA06459.1; PID
 A:Note: parts of this sequence, including the amino ends of three chains isolated from t
 C:Comment: The mechanism of association with the membrane of the intestinal brush border
 C:located below) or with amino-terminal myristoylation of the heavy chain.
 C:Complex: Mature enteropeptidase is variously reported to contain two (heavy and light)
 ed by a disulfide bond. Possibly, conversion from membrane-bound to soluble forms involv
 C:Function:

A:Description: Cleaves activation peptide from trypsinogen to produce active trypsin
 A:Pathway: intestinal digestive hydrolase cascade
 C:Superfamily: enteropeptidase; C1r/C1s repeat homology; LDL receptor ligand-binding rep
 C:Keywords: glycoprotein; hydrolase; serine proteinase; transmembrane protein; zymogen
 F:22-38/Domain: transmembrane #status predicted <TM>
 F:52-117/Product: enteropeptidase mini chain #status predicted <MCH>
 F:118-799/Product: enteropeptidase heavy chain #status predicted <HCH>
 F:119-236/Domain: LDL receptor ligand-binding repeat homology <LDL1>
 F:357-519/Domain: MAM homology <MAM>
 F:541-646/Domain: C1r/C1s repeat homology <C1R>
 F:658-692/Domain: LDL receptor ligand-binding repeat homology <LDL2>
 F:693-798/Domain: scavenger receptor cysteine-rich domain homology #status atypical <SRC
 F:800-1034/Product: enteropeptidase light chain #status predicted <LCH>
 F:800-1029/Domain: trypsin homology <TRY>
 F:116,147,170,194,283,343,350,403,455,485,518,549,645,697,701,721,740,761,804,863,902,96
 F:787-911,925-992,956-971,982-1010/Disulfide bonds: #status predicted
 F:840,891,986/Active site: His, Asp, Ser #status predicted

Query Match 23.6%; Score 552; DB 1; Length 1034;
 Best Local Similarity 34.2%; Pred. No. 1.7e-37;
 Matches 127; Conservative 69; Mismatches 141; Indels 34; Gaps 14;

```

QY 74 LDCPLGDEDEHCVKSPPEGPAAVRLSKDRSTLQVLDATGWMFSAFCPNFTEAL 133
DB 676 LDCGFSHCKGSDDAHCVR-PLNGTANNGSLVQFR--IQSI-----MHWACAEWTTQT 726
QY 134 AETACRQMGY-----SSKPTFAVEIGPDODLVEITENSQELMRNMSGCLSGSLVSL 189
DB 727 SDDVQCQLLGLGTGNSMNPFFSSG-GP-----FKIMLTAPGSLILTNASBQCFPDSJLL 780

```


QY 354 CNADADAYCGEVLTEKXMKCGIPEGGVDTCCGDSGGPLMYO-----SPDMHVLVIVSNGYGC 408

Db 322 CNGADFLYNGIQRPKVFCAGIPEGGIIACQDSSGFLVCEDSIRTRWRKLCGLIVSNCTGC 381

QY 409 GGEPTGVTTKVSAYLNMIVYNWK 432

Db 382 ALAQKPGVYTKVSDFERWILFQAIK 405

RESULT 2

enteropeptidase (EC 3.4.21.9) precursor [validated] - human
N/Alternate names: enteroxinase
C/Species: Homo sapiens (man)
C/Date: 19-May-1995 #sequence_revision 09-Aug-1996 #text_change 09-Jul-2004
C/Accession: A56318; B43090
R/Kitamoto, Y.; Veale, R.A.; Donis-Keller, H.; Sadler, J.E.
Biochemistry 34, 4562-4568, 1995
A/Title: cDNA sequence and chromosomal localization of human enteroxinase, the proteolytic
A/Reference number: A56318; MUID:95234679; PMID:7718557
A/Accession: A56318
A/Molecule type: mRNA
A/Residues: 1-1019 <KIT>
A/Cross-references: UNIPROT:P98073; UNIPARC:UPI000003FE65; GB:U09860; NID:G746412; PIDD:
R.Kitamoto, Y.; Yuan, X.; Wu, Q.; McCourt, D.W.; Sadler, J.E.
Proc. Natl. Acad. Sci. U.S.A. 91, 7588-7592, 1994
A/Title: Enteroxinase, the initiator of intestinal digestion, is a mosaic protease compo
A/Reference number: A43090; MUID:94329561; PMID:8052624
A/Accession: B43090
A/Status: nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 749-1019 <KIT2>
A/Cross-references: UNIPARC:UPI0000172B0B; GB:U09860
C/Comment: The mechanism of association with the membrane of the intestinal brush border
created below) or with amino-terminal myristoylation of the heavy chain.
C/Genetics:
A/Gene: GDB:PR857
A/Cross-references: GDB:384083; OMIM:226200
A/Map position: 21q21-21q21
C/Complex: Mature enteropeptidase is variously reported to contain two (heavy and light)
ed by a disulfide bond. Possibly, conversion from membrane-bound to soluble forms involv
ducts.
C/Function:
A/Description: cleaves activation peptide from trypsinogen to produce active trypsin
A/Pathway: intestinal digestive hydrolase cascade
C/Superfamily: enteropeptidase; C1r/C1s repeat homology; LDL receptor ligand-binding reg
C/Keywords: glycoprotein; hydrolase; serine proteinase; transmembrane protein; zymogen
F.1-784/Product: enteropeptidase heavy chain #status predicted <HCH>
F.122-18/Domain: transmembrane #status predicted <TM>
F.184-221/Domain: LDL receptor ligand-binding repeat homology <LDL1>
F.342-504/Domain: MAM homology <MAM>
F.526-631/Domain: C1r/C1s repeat homology <C1R>
F.643-677/Domain: LDL receptor ligand-binding repeat homology <LDL2>
F.678-783/Domain: scavenger receptor cysteine-rich domain homology #status atypical <SRC
F.785-1019/Product: enteropeptidase light chain #status predicted <LCH>
F.785-1014/Domain: trypsin homology <TRY>
F.116,147,179,328,335,388,440,470,503,534,630,682,706,725,848,887,909,949/Binding site:
F.1772-896,810-826,910-977,941-956,967-995/Disulfide bonds: #status predicted
F.1825,876,971/Active site: His, Asp, Ser #status predicted

	Query Match	24.9%	Score 582.5	DB 1	Length 1019
	Best Local Similarity	33.8%	Pred. No. 56-40		
	Matches 135	Conservative 62	Mismatches 132	Indels 71	Gaps 15
Qy	67 HF-----IPKQICDGLDPCLDGDEHCVKSPFEGRAVAVRLSKRSTIQVLDSATG	119			
Db	647 HFQCKNGCEVPLVNLCDGHLHCEGSDDEADCVRF-----NGTTN	686			
Qy	120 N-----WFSACFDNFTTEALAEACRQWY-----SSKPTRAVEIGDQDLDVE	164			
Db	687 NNGLVFRFIQSIWTHCAENWTTQISNDVQCQLGLGSGSSKPIF-STDGGFVFKLNTAP	745			
Qy	165 -----ITENSOELRMRRSSGPECLSGSLVSLYHLC--LACGSKLK-----TPRVGGEAASVDSW	214			

Db	746	DGHLILTPSSQ-----CLQDSLIRLQCHNHKCGCKLLAODITPKLVGSSNALEGAW	796
Qy	215	FMQVSIQYDKOHVCGSILDPHVVLTAHC-FRKHDTFNMKVRGSDKLGSEFPIAVAK	273
Db	797	FMVVLGYVGGRIILCASLVSDDMLVSAHCYVGRNLPEBSKPAIIGLHNKSNLITSQITVP	856
Qy	274	IIII--IEENPMY---PKNDILALMKLOPELTFSGVIRPILCFEFDDELTPATPLWIIIGW	328
Db	857	RLIDEIVNPHNRRKNDIMMHELFVANTDYIQPCLDEENOVPPGRGCSIAIGW	916
Qy	329	FTKONGGKMSDILQASVQVIDSTRCNMDADVQGEVTEPMQACGIEGGVDTCQSGSGP	388
Db	917	-TVVYQGTIANLQADVPLFLENKQ--QOMEYNITENMICAGYEGGIDSCQSGSGP	974
Qy	389	LMYQ--SDQMHVVGIVSWGCGCGPSTPGVYTYTSAYLANMI	427
Db	975	LMQDENNRWFLAGVTSFGKCALPNRPGYAYAVSRTEMI	1014

RESULT 3

A:3030
 enteropeptidase (EC 3.4.21.9) precursor [validated] - bovine
 N:Alternate names: enterokinase
 C:Species: Bos primigenius taurus (catle)
 C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 09-Jul-2004
 C:Accession: A43090; A46874; A61436
 R:Kiltemoco, Y.; Yuan, X.; Wu, Q.; McCourt, D.W.; Sadler, J.E.
 Proc. Natl. Acad. Sci. U.S.A. 91, 7588-7592, 1994
 A:Title: Enterokinase, the initiator of intestinal digestion, is a mosaic protease composed of two distinct domains
 A:Reference number: A43090; MWID:94329561; PMID:8052624
 A:Accession: A43090
 A>Status: nucleic acid sequence not shown; translated from GB/EMBL/DBD
 A:Molecule type: mRNA
 A:Residues: 1-1035 <KT>
 A:Cross-references: UNIPROT:P98072; UNIPARC:UPI000004BBB5; GB:U09859; NID:9746410; PID:NID:9746410
 A:Experimental source: small intestine
 R:Valle, E.R.; Rehentulla, A.; Reile, L.A.; DiBlasio, E.A.; Ferez, C.; Grant, K.L.; et al.
 J. Biol. Chem. 268, 23311-23317, 1993
 A:Title: Cloning and functional expression of a cDNA encoding the catalytic subunit of bovine enterokinase
 A:Reference number: A48874; MWID:94043122; PMID:8226855
 A:Accession: A48874
 A:Molecule type: mRNA
 A:Residues: 801-1035 <LAV>
 A:Cross-references: UNIPARC:UPI000011133D; GB:U19663; NID:9416131; PID:AAA16035.1; PID:519663
 A:Note: parts of this sequence, including the amino end of the mature protein, were confirmed by sequencing
 R:Light, A.; Janska, H.
 J. Protein Chem. 10, 475-480, 1991
 A:Title: The amino-terminal sequence of the catalytic subunit of bovine enterokinase.
 A:Reference number: A61436; MWID:92189715; PMID:1799406
 A:Accession: A61436
 A:Molecule type: protein
 A:Residues: 801-807, 'Y', 809-827 <LIG>
 A:Cross-references: UNIPARC:UPI00001468A0
 C:Comment: The mechanism of association with the membrane of the intestinal brush border
 C:Comment: The mechanism of association with the membrane of the intestinal brush border
 C:Comment: Conversion from membrane-bound to soluble forms may involve further processing
 C:Complex: mature enteropeptidase is variously reported to contain two (heavy and light) subunits
 I:Id: linked
 C:Function:
 A:Description: cleaves propeptide from trypsinogen to produce active trypsin
 A:Pathway: intestinal digestive hydrolase cascade
 C:Superfamily: enteropeptidase; Ctr/Cls repeat homology; LDL receptor ligand-binding repeat homology
 C:Keywords: glycoprotein; hydrolase; intestine; serine proteinase; transmembrane protein
 F:22-38/Domain: transmembrane #status predicted <TM>
 F:52-117/Product: enteropeptidase mini chain #status predicted <MC>
 F:118-800/Product: enteropeptidase heavy chain #status predicted <HC>
 F:199-236/Domain: LDL receptor ligand-binding repeat homology <LDL1>
 F:358-520/Domain: MAM homology <MAM>
 F:542-647/Domain: Ctr/Cls repeat homology <CLR>
 F:559-693/Domain: LDL receptor ligand-binding repeat homology <LDL2>
 F:694-1093/Domain: scavenger receptor cysteine-rich domain homology #status atypical <SRC>
 F:694-1093/Product: enteropeptidase light chain #status predicted <LC>
 F:694-1093/Product: enteropeptidase light chain #status predicted <LC>

353 ILSKVQASVLELIDCISVITK

THIS PAGE BLANK (USPTO)

Qy 405 GYCGGPGSTPGVYTKVSAIYLMWYVWK 432
Db 230 GIGCARPKPKPGVYVRSFVWTIEGVNR 257

RESULT 12

US-11-056-621-4
; Sequence 4, Application US/11056621
; Publication No. US20050262592A1
; GENERAL INFORMATION:
; APPLICANT: Spencer, David
; APPLICANT: Dickey, Lynn F.
; APPLICANT: Gaadaska, John R.
; APPLICANT: Wang, Xiaowei
; APPLICANT: Cox, Kevin M.
; APPLICANT: Peele, Charles G.
; TITLE OF INVENTION: EXPRESSION OF PLASMINOGEN AND
; TITLE OF INVENTION: MICROPLASMINOGEN IN DUCKWEED
; FILE REFERENCE: 40989/274646
; CURRENT APPLICATION NUMBER: US/11/056,621
; PRIOR FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: US 60/543,487
; PRIOR FILING DATE: 2004-02-11
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 791
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Sequence of mature human plasminogen
US-11-056-621-4

Query Match 18.7%; Score 438.5; DB 7; Length 791;
Best Local Similarity 37.6%; Pred. No. 2.3e-36;
Matches 94; Conservative 43; Mismatches 94; Indels 19; Gaps 6;

Qy 194 CGKSLKTP----RVGGEESVDSWPMQVSIQVD-KQHVCGGSLDHPHWLTAHCRK 247
Db 548 CGKQVBEKPKCPGRVGGCVAHPSHPQVSLRTFGHFCGCTLISPEWVLTAAHCLKR 607
Qy 248 HTDVFNMKVRAGSDLSF----PSLAVALKIIIEFNPMYPKNDIAIMKQFPLTFSGT 303
Db 608 SPRBSYVVLGAHQEVLEPHVQIEVSRLE----LEPRKDIALLKLSPPAVITDK 661
Qy 304 VRPCLPFDEBELPATPLMIIGWFTKONGKMSDILLQASVOYIDSTRCNADDAVYGE 363
Db 662 VIPACLPEPNVAVADRTCEFTVGEGTGTG--AGLKEAOLPIENKVCNRYEFLNGR 719
Qy 364 VTEKMMCAIGPEGVDTCQDGGGGLM-YQSDQMHVGIIVSWGCGGSPSTPGVYTKVSA 422
Db 720 VQSTELCAGHLAAGTDSQDGGGGLPVCPEKDKYLLQGVTSGLGACARPNKRGVYVRSR 779
Qy 423 YLMWYVWK 432
Db 780 FVTWIEGVNR 789

RESULT 13

US-10-131-826A-506
; Sequence 506, Application US/10131826A
; Publication No. US20050245730A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Bereini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.

; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C128
; CURRENT APPLICATION NUMBER: US/10/131,826A
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 506
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-131-826A-506

Query Match 15.9%; Score 371.5; DB 6; Length 250;
Best Local Similarity 35.2%; Pred. No. 2.5e-30;
Matches 90; Conservative 40; Mismatches 99; Indels 27; Gaps 9;

Qy 189 LHCLACGSKLTPRVGGEESVDSWPMQVSIQVDKQHVCGGSLDHPHWLTAHCRK- 247
Db 8 LLAATLGVGETETIINGFECKPSQPMQALFEKTRLICATLILAPRWLTLTAHCLKPR 67
Qy 248 ---HTDVFNMKVRAGSDK---LGSFSLAVAKIIIEFNPMY-KD--NDIAIMKQFPL 297
Db 68 YIVHLGQHNLQKEGCEQRTATSFPHPG-----FNNSLPNKHNDIMLVKMASP 119
Qy 298 LTFSGYVRPCLPFDEBELPATPLMIIGWFTKONGKMSDILLQASVOYIDSTRCNAD 357
Db 120 VSIWAVRPLTL--SRCVTAAGTSCILISGWSTSSPOLRLPHTLRCAVITLIEHOKC--E 175
Qy 358 DAYQGEVTEKMMCAIGPEGVDTCQDGGGGLMAYQSDQMHVGIIVSWGCGGSPSTPGV 416
Db 176 NAYFGNITDMVCAISVQEGGSDSCQDGGGGLPVNGS---LQGISNWDCAITRRKGV 232
Qy 417 YTKVSAIYLMWYVWK 432
Db 233 YTKVCKVDMWQETMK 248

RESULT 14

US-11-029-003-6
; Sequence 6, Application US/11029003
; Publication No. US20050260194A1
; GENERAL INFORMATION:
; APPLICANT: PETERS, ROBERT T.
; APPLICANT: MEZO, ADAM R.
; APPLICANT: RIVERA, DANIEL S.

```

; GENERAL INFORMATION:
; APPLICANT: Ma, Jian-xing
; TITLE OF INVENTION: COMPOUNDS USEFUL IN INHIBITING VASCULAR LEAKAGE, INFLAMMATION AND
; FILE REFERENCE: 5820, 656
; CURRENT APPLICATION NUMBER: US/11/010,874
; PRIOR FILING DATE: 2004-12-11
; PRIOR APPLICATION NUMBER: 10/963,115
; PRIOR FILING DATE: 2004-10-12
; PRIOR APPLICATION NUMBER: 60/510,620
; PRIOR FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: 60/528,647
; PRIOR FILING DATE: 2003-12-11
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
; LENGTH: 812
; TYPE: PRT
; ORGANISM: murine
US-11-010-874-1
```

```

Query Match          19.4%; Score 455.5; DB 7; Length 812;
Best Local Similarity 39.4%; Pred. No. 4.6e-38;
Matches 102; Conservative 43; Mismatches 89; Indels 25; Gaps 8;
```

```

QY 181 CLEGSIVSLHCLACGKSLKTP-----RVGGEASVDSWPMQVSI--QYDKOHVCGSITL 233
DB 560 CASAS-----SFECKRQVEPKKCPGRVVGCVANPHSMWQSLKRTFGQFCGTLI 614
QY 234 DPHWVLTAAHCFKRDVFNWVKYRAGSDK--IG-SFPSIAVAKIIEFNPMYPRKNDI 289
DB 615 APFWVLTAAHCEKSRPEFYKYLGAHEVIRGLDQVETSAKLI-----LEPNRDI 668
QY 290 ALMKLOPFLFSGTVRRCICPFDEBELTPATPLMIITGWETKONGKMSDILIQASVOYI 349
DB 669 ALMKLSRPAITITKVPACLPSPNYVADRTICYITGWETQGTFG--AGRLKEAQLPVI 726
QY 350 DSTRCNADAVQGEVTEKMKACGIPREGVDTCCGDSGGPLM--YQSDOMHVIVGIVSV 408
DB 727 ENKVCNRVEYLNRRVSTELCAQQLAGVDSCCGDSGGPLVCEKXKYLILQGVTSWGLCC 786
QY 409 GGPSTGGVTVTKYSAVYLNWI 427
DB 787 ARPNKPGVYVVRSPYDWI 805
```

```

RESULT 10
US-10-821-234-1043
; Sequence 1043, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: PL SEQ_genes Version 1.0
; SEQ ID NO 1043
; LENGTH: 393
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1043
```

```

Query Match          19.2%; Score 449; DB 6; Length 393;
Best Local Similarity 36.6%; Pred. No. 7.6e-38;
Matches 113; Conservative 40; Mismatches 116; Indels 40; Gaps 10;
```

```

QY 147 PTFRAVEI--GPPQDLVVEITENSQELMRNSGCPCLSGSLVSLHCLA----- 193
DB 35 PTFRAVSLQGP-----LSWANAQKVLDGGQLGAVALILYGLRSGTGAEGA 83
QY 194 ---CGKSLTPRVVGGEEASVDSWPMQVSIQYDKOHVCGSITLDPHWVLTAAHCFKRD 250
DB 84 EAPCGVAPQA-RITGSSSAVAGQWPMQVSIYEGVHVGGSIVSQWVLTAAHCFKRD 142
QY 251 VFNWVKRAGSDKLGSEFSLA-VAKIIEFNPMYPRK--NDIALMKLOPFLFSGTVR 306
DB 143 KEAYEVKGAHQLDSEYSEDAKYSTIKDIIHPYSYLOEGSQGDIALQLSRPTTFRYIRP 202
QY 307 ICLPFDEBELTPATPLMIITGWETKONGKMS-DILQASVOYIDSTCMA---DDAQG 362
DB 203 ICLPAAVASFPNGHCTVYTGWGHVAPSVSLTPKPLQOEVLISRETQCLYNIDAKE 262
QY 363 E---VTEKMKACGIPREGVDTCCGDSGGPLMYQSD-QMHWVIVGVMGCGGSPSTPGYTT 418
DB 263 EPHFVQEDWVCAGYVEGGKDAQCQDSGGPLSCPVGLMYLIGIVSGDCAKARNRPYTT 322
QY 419 KVSAYLNWI 427
DB 323 IASSYASWT 331
```

```

RESULT 11
US-11-056-621-6
; Sequence 6, Application US/11056621
; Publication No. US20050262592A1
; GENERAL INFORMATION:
; APPLICANT: Spencer, David
; APPLICANT: Dickey, Lynn F.
; APPLICANT: Gasdaska, John R.
; APPLICANT: Wang, Xiaowei
; APPLICANT: Cox, Kevin M.
; APPLICANT: Peele, Charles G.
; TITLE OF INVENTION: EXPRESSION OF PLASMINOGEN AND
; FILE REFERENCE: 40989/274646
; CURRENT APPLICATION NUMBER: US/11/056,621
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: US 60/543,487
; PRIOR FILING DATE: 2004-02-11
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Sequence of mature human microplasmnogen
US-11-056-621-6
```

```

Query Match          18.9%; Score 442.5; DB 7; Length 259;
Best Local Similarity 36.2%; Pred. No. 1.9e-37;
Matches 97; Conservative 44; Mismatches 98; Indels 29; Gaps 7;
```

```

QY 176 NSSGPCLSGSLVSLHCLACGKSLKTP-----RVGGEASVDSWPMQVSIQYD-KOHVCG 229
DB 8 NTQGPSTP-----CGKQVEPKKCPGRVVGCVANPHSMWQSLKRTFGMFCG 57
QY 230 GSILDPHWVLTAAHCFKRDVFNWVKYRAGSDKLSF---PSIAVAKIIEFNPMYPRK 285
DB 58 GTLISPFWVLTAAHCEKSRPEFSYKYLGAHQGVNLEPHVQIEVERLP-----LEPT 111
QY 286 DNDIALMKLOPFLFSGTVRRCICPFDEBELTPATPLMIITGWETKONGKMSDILIQAS 345
DB 112 RKDIALKLSPPAVITDKVPACLPSPNYVADRTICYITGWETQGTFG--AGRLKEAQ 169
QY 346 VQVIDSTRCNADAVQGEVTEKMKACGIPREGVDTCCGDSGGPLM--YQSDOMHVIVGIVSV 404
DB 170 LPVIENTVCNRYEFTLNRRVOSTELCAQHLAGVDSCCGDSGGPLVCEKXKYLILQGVTSW 229
```

Db 413 LRDMI 417

RESULT 7

US-11-137-465-56

; Sequence 56, Application US/11137465
; Publication No. US20050255558A1

; GENERAL INFORMATION:

; APPLICANT: Agarwal, Pankaj
; APPLICANT: Murdock, Paul R.
; APPLICANT: Rizvi, Safia K.
; APPLICANT: Smith, Randall, F.
; APPLICANT: Xiang, Zhaoying
; APPLICANT: Kadnick, Karen

; TITLE OF INVENTION: NOVEL COMPOUNDS

; FILE REFERENCE: Gp50018

; CURRENT APPLICATION NUMBER: US/11/137,465

; CURRENT FILING DATE: 2005-05-25

; PRIOR APPLICATION NUMBER: US/10/239,663

; PRIOR FILING DATE: 2002-09-24

; PRIOR APPLICATION NUMBER: PCT/US01/09226

; PRIOR FILING DATE: 2001-03-22

; PRIOR APPLICATION NUMBER: 60/192,158

; PRIOR FILING DATE: 2000-03-24

; PRIOR APPLICATION NUMBER: 60/192,668

; PRIOR FILING DATE: 2000-03-27

; PRIOR APPLICATION NUMBER: 60/200,166

; PRIOR FILING DATE: 2000-04-27

; NUMBER OF SEQ ID NOS: 66

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 56

; LENGTH: 305

; TYPE: PRT

; ORGANISM: Homo sapiens

US-11-137-465-56

Query Match 20.2%; Score 473.5; DB 7; Length 305;
Best Local Similarity 37.8%; Pred. No. 1.8e-40;

Matches 95; Conservative 43; Mismatches 100; Indels 13; Gaps 4;

QY 193 ACGSKLTPRVVGEASVDSMPQVSIQYDKOHVCGSILDPHVLTLAAHCFRRHTDVF 252

Db 32 ACGREHIALVAGVESARGRPWQASLRRLRRHRCGSLLSRRVLSAAHCFQHYPS 91

QY 253 NMKTRAGS-----DKLGFPSLAVAKIIIEFNPMYPKNDIALMLQPLTFESGTVR 305

Db 92 EMTQVLGSLTRPFRPMNLRAVSSRKVQDIIIVPALGLVANDIALRLASSVTYNATIQ 151

QY 306 PICLPFDEELTPATPLMIIGWFTKONGKMSD--ILLQASVOYIDSTRCN--ADDAV 360

Db 152 PICTESSTFNVRHDDCVTGMGLISPGTPLRPYYNLREAQVTLINNTRCNYLFEQPS 211

QY 361 QGEVTEKMKAGIPEGGVDTCCGDSGGPLMTQSD-QMHVGVISKVGCCGGSTPGVYTK 419

Db 212 RSMWDSMFCAEGADGSDVDTCKDGGSLVCDKGLMYQVGVISWMDCCGPNRPGVYTN 271

QY 420 VSAVLMNLYNV 430

Db 272 ISVTFHMTIRRV 282

RESULT 8

US-10-131-826A-222

; Sequence 222, Application US/10131826A

; Publication No. US20050245730A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Beresini, Maureen

; APPLICANT: Deforge, Laura

; APPLICANT: Demoyers, Luc

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Sherwood, Steven

; APPLICANT: Smith, Victoria

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tamas, Daniel

; APPLICANT: Watanabe, Colin K

; APPLICANT: Wood, William

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P33081C128

; CURRENT APPLICATION NUMBER: US/10/131,826A

; CURRENT FILING DATE: 2002-04-24

; PRIOR APPLICATION NUMBER: 60/049911

; PRIOR FILING DATE: 1997-06-18

; PRIOR APPLICATION NUMBER: 60/056974

; PRIOR FILING DATE: 1997-08-26

; PRIOR APPLICATION NUMBER: 60/059113

; PRIOR FILING DATE: 1997-09-17

; PRIOR APPLICATION NUMBER: 60/059115

; PRIOR FILING DATE: 1997-09-17

; PRIOR APPLICATION NUMBER: 60/059117

; PRIOR FILING DATE: 1997-09-17

; PRIOR APPLICATION NUMBER: 60/059122

; PRIOR FILING DATE: 1997-09-17

; PRIOR APPLICATION NUMBER: 60/059184

; PRIOR FILING DATE: 1997-09-17

; PRIOR APPLICATION NUMBER: 60/059263

; PRIOR FILING DATE: 1997-09-18

; PRIOR APPLICATION NUMBER: 60/059352

; PRIOR FILING DATE: 1997-09-19

; PRIOR APPLICATION NUMBER: 60/059388

; PRIOR FILING DATE: 1997-09-19

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 550

; SEQ ID NO 222

; LENGTH: 290

; TYPE: PRT

; ORGANISM: Homo Sapien

US-10-131-826A-222

Query Match 20.0%; Score 467.5; DB 6; Length 290;
Best Local Similarity 37.1%; Pred. No. 6.9e-40;

Matches 99; Conservative 49; Mismatches 92; Indels 27; Gaps 9;

QY 184 GSLVSLHCLACGSKLTPRVVGEASVDSNPQVSIQYDKOHVCGSILDPHVLTLAAH 243

Db 16 GSGRAKATACGRPMNRWVGQDTQEGEWPQVSIQKNGSHFCGSLIAEQVLTAAH 75

QY 244 CFRKHTDVFNKTVAGSDKL-GSFPSLAVAKIIIEFNPMY---PKNDIALMLQPLT 299

Db 76 CFRNTSSTLSYQVLGLRQLVQPPHMYAVRQVENSPLTQTAASADVALVLEAVP 135

QY 300 FSGTVRPLCLPFPEBELTPATPL--WIIIGWFTKONGKMSD-----ILLQASVOYIDS 351

Db 136 FTNYILPVCUP--DPSVIFETGMNCWVTGWSPE-----EDLLPPEITLKLAVPIIDT 188

QY 352 TRCN-----ADDAVQGE-VTEKMKAGIPEGGVDTCCGDSGGPLMTQSDQ-WHVGVIS 403

Db 189 PKCNLYSKDTEFPYQPKTIKNDLACGFEKGDACKGDSGGPLVCLVQGSWILQAGVIS 248

QY 404 WGYCGGSGSTPGVYTKYSAYLMNLYNV 430

Db 249 WGEBCARQNRPGYIVRVTAHNMWTHRI 275

RESULT 9

US-11-010-874-1

; Sequence 1, Application US/11010874

; Publication No. US20050250694A1

; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4
; LENGTH: 1113
; TYPE: PRT
; ORGANISM: Mus Musculus
US-11-067-811-4

Query Match 20.9%; Score 488.5; DB 7; Length 1113;
Best Local Similarity 33.2%; Pred. No. 3.4e-41;
Matches 130; Conservative 52; Mismatches 148; Indels 61; Gaps 17;

QY 69 IPRKOLCDGELDPLGEDEEHCVKSPFEGPAAVAVRLSKDRSTLQVLDATGMMFSAQCFPN 128
DB 736 VRDLMCDGMDVDCSSSDSDGCVTLSSKNG-----NSSSLTLVHKSAREH--HYCADG 785
QY 129 FTEALAEATACROMGYSSKRTFRAVEIGPRODDLVETINSQELRN---RMSG----- 179
DB 786 WRETLSQLACKQMGJGEPSTVKLI--PQO-----DGQOWLRLYPMWENLNGSTLOE 834
QY 180 -----PCLSGSLVSLHCLA--CGK---SLKTPRVVGESEASVDSWPMQVSIQYDKQ-HV 227
DB 835 LLVYHNSGCSRSREISLCLKQDCGRPARAAMKRILIGRTSRGRMPQCSLQSEPSGHI 894
QY 228 CGSILDPHMLVLAHCPEKRTDVFNMKYRAGSDKL--GSPFSLAVAKIIIEFNPMYP 284
DB 895 CGCVLIARKKVLTVAHCFEGREDADVMKVFGINNLDHPSGFQTRFVKTILL--HPRYS 952
QY 285 K---DNDILAMKLOPFLTSGTVRPICLPFDEBLTPATPLMIGMFTKONGKMSDIL 341
DB 953 RAVVYDIDSVBELSDINETSIVRPVCLPSPEEYLEPDYCYITGWC--HMGNKMPFKL 1009
QY 342 LQASVQIDSTCNADAY--QGEVTEKMMKACGIPREGVDTCCGDSGGPLMYQ--SDQW 397
DB 1010 QGEVRIITPLEQC--QSYFDMKTIITRMICAGYESGTVDSMGDSGGLVCEPREGQNT 1066
QY 398 VVGIVSMGYGCGGPGST-PGYTTKVSAYLWMI 427
DB 1067 LFGLTSMGSGVCFPSKVLGPGYVSNVSFYEWMI 1097

RESULT 5
US-11-067-811-1
; Sequence 1, Application US/11067811
; Publication No. US20050260688A1
; GENERAL INFORMATION:
; APPLICANT: Morgan, Bruce A
; APPLICANT: Enshel1-SciJffers, David
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR HAIR GROWTH
; FILE REFERENCE: 10287-083001
; CURRENT APPLICATION NUMBER: US/11/067,811
; CURRENT FILING DATE: 2005-02-28
; PRIOR APPLICATION NUMBER: US 60/548,272
; PRIOR FILING DATE: 2004-02-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
; LENGTH: 1042
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-067-811-1

Query Match 20.7%; Score 484; DB 7; Length 1042;
Best Local Similarity 32.6%; Pred. No. 8.9e-41;
Matches 126; Conservative 64; Mismatches 146; Indels 50; Gaps 17;

QY 69 IPRKOLCDGELDPLGEDEEHCVKSPFEGPAAVAVRLSKDRSTLQVLDATGMMFSAQCFPN 128
DB 666 VRDLMCDGMDVDCSSSDSDGCVTLSSKNG-----TISINNS-SFLMVAATEH--HYCADG 717
QY 129 FTEALAEATACROMGYSSKRTFRAVEIGPRODDLVETINSQELRN---RMSG----- 178
DB 718 WRETLSQLACKQMGJGEPSTVKLIQ---DQEKPRMVLTHSNWESLNGTTLHELLVNGGS 774

QY 179 GPCLSGSLVSLHCLA--CGK---SLKTPRVVGESEASVDSWPMQVSIQYDKQ-HVCGGSI 232
DB 775 --CESRSKISLILCTKQDCGRPARAAMKRILIGRTSRGRMPQCSLQSEPSGHI 832
QY 233 LDPHMLVLAHCPEKRTDVFNMKYRAGSDKL--GSPFSLAVAKIIIEFNPMYPK--D 286
DB 833 IAKKVVLTVAHCFEGRENAAVWKVLAGINNLDHPSVFMQTRPVKTIIL--HPRYSRAVD 890
QY 287 NDIALMKLOPFLTSGTVRPICLPFDEBLTPATPLMIGMFTKONGKMSDILQASV 346
DB 891 YDISIVELSEDISESGTVRPICLPPEQMLEPDYCYITGWC--HMGNKMPFKLOEGEV 947
QY 347 QVIDSTRCNADAY--QGEVTEKMMKACGIPREGVDTCCGDSGGPLMYQ--SDQWVAV 402
DB 948 RILSLHLC--QSYFDMKTIITRMICAGYESGTVDSMGDSGGLVCEKPGGRMTLFLGLT 1004
QY 403 SMGYCGGPGST-PGYTTKVSAYLWMI 427
DB 1005 SMGSGVCFPSKVLGPGYVSNVSFYEWMI 1030

RESULT 6
US-11-102-240-106
; Sequence 106, Application US/1102240
; Publication No. US20050260647A1
; GENERAL INFORMATION:
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: ANTIBODIES TO POLYPEPTIDES ENCODED BY A NUCLEIC ACID UNDEREXPRESS
; FILE REFERENCE: P3230R1C106C
; CURRENT APPLICATION NUMBER: US/11/102,240
; CURRENT FILING DATE: 2005-04-08
; PRIOR APPLICATION NUMBER: 10/063662
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 10/006867
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 60/170262
; PRIOR FILING DATE: 199-12-09
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 106
; LENGTH: 423
; TYPE: PRT
; ORGANISM: Homo Sapien
US-11-102-240-106

Query Match 20.3%; Score 475; DB 7; Length 423;
Best Local Similarity 39.6%; Pred. No. 2.1e-40;
Matches 97; Conservative 41; Mismatches 97; Indels 10; Gaps 5;

QY 190 HCLAG----KSL-KTPRVVGESEASVDSWPMQVSIQYDKQHVCGGSLDHPHMLVLAHC 244
DB 176 HC--CGTRRSKTLGSLRIVGSTVEEGEEMPOQASLDWDSHRCATTLINATWLVSAHC 233
QY 245 FRKHTDVFNMKYRAGSDKLGSFSLAVAKIIIEFNPMYPKNDIALMKLOPFLTSGTV 304
DB 234 FTTYNPAPRWASFEVITIKPSMKRGLRIIVHEKXKPSHDYDLSLAEISSPYVTNAV 293
QY 305 RPICLPFDEBLTPATPLMIGMFTKONGKMSDILQASVQVIDSTRCNADAYQGEV 364
DB 294 HNVCLPDASYEPQPDVNFVTGFGALK-NDGYSQWHLQAOVTLIDATTCNBPQAYNDAI 352
QY 365 TEKMMKACGIPREGVDTCCGDSGGPLMYOS--DQWVAVIGVSMGYGCGGPGSTPGYTTKYS 422
DB 353 TPRMLCAASLEBKTDACOGDSGGLVSSDARDIWTLAGIWSGDECAKPNKPGVYTRVYA 412
QY 423 YLWMI 427

Db 179 LSGSLVSLHCLACGSKLTPRVVGGEEASVDSWPMQVSIQYDKQHVCGSILDPHWLTA 238
Qy 242 AHCFRKHDTVDYFNKVRGSDKLSFPSLAVAKIIIEFNMPYKNDIALMKLOPFLFS 301
Db 239 AHCFRKHDTVDYFNKVRGSDKLSFPSLAVAKIIIEFNMPYKNDIALMKLOPFLFS 298
Qy 302 GYRPICLPFDEBELTPATPLMIIGWFTKONGKMSDILLQASVQYIDSTRCNADDAAY 361
Db 299 GYRPICLPFDEBELTPATPLMIIGWFTKONGKMSDILLQASVQYIDSTRCNADDAAY 358
Qy 362 GEVTEKMACAGIPREGVDTCCGDSGGPLMTQSDQMHVGVISWGYCGGPGSTPGVTVKVS 421
Db 359 GEVTEKMACAGIPREGVDTCCGDSGGPLMTQSDQMHVGVISWGYCGGPGSTPGVTVKVS 418
Qy 422 AYLNMTYVWKAKEL 435
Db 419 AYLNMTYVWKAKEL 432

RESULT 2
US-11-182-752-4
; Sequence 4, Application US/11182752
; Publication No. US20050250154A1
; GENERAL INFORMATION:
; APPLICANT: Gan, Weiniu et al.
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: CL001194DIVII
; CURRENT APPLICATION NUMBER: US/11/182,752
; CURRENT FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 417
; TYPE: PRT
; ORGANISM: HUMAN
US-11-182-752-4

Query Match 25.1%; Score 588.5; DB 7; Length 417;
Best Local Similarity 31.5%; Pred. No. 7.7e-52;
Matches 140; Conservative 60; Mismatches 161; Indels 83; Gaps 11;

Qy 19 KPRIMEFRKVGIPITIALSLASITIVVLIKVIDKYFLCGQPLHFPKQQLDGE 78
Db 15 RPKVAALT---AGTLILLTAIGASMAIVAVLLR----- 45
Qy 79 LDCPLGDEDEHCVKSPFEGPAVAVRLSKDRSTLOVLDATGNWFSACFDNFTALAEATAC 138
Db 46 -----SDQE-----PLYPVQVSSADARLMVFDKTEGTWRLLCSSRSNARVAGLSC 90
Qy 139 ROMGYSSKPTFAVEIGPDOLDVETEN-----SOELRMKNSGPGCLS 183
Db 91 EEMGF-----LRAL---THSELDTVATAGNIGSFVFCVDEBGLPHTQHLLEVITSPDCPR 142
Qy 184 GSLVSLHCLACG-KSLKTPRVVGGEEASVDSWPMQVSIQYDKQHVCGSILDPHWLTA 242
Db 143 GRFLAICODCGRRKLPVDRIVGGRDTSIGRPMQVSLRYDGAHLCGSLLSGDWVLTAA 202
Qy 243 HCFRKHDTVDYFNKVRGSDKLSFPSLAVAKIIIEFNMPY-----KNDIALMK 293
Db 203 HCFRKHDTVDYFNKVRGSDKLSFPSLAVAKIIIEFNMPY-----KNDIALMK 262
Qy 294 LQFPLTFSGTVAPICLPFDEBELTPATPLMIIGWFTKONGKMSDILLQASVQYIDSTR 353
Db 263 LSSPLRLEYIQPVCLPAAGQALVVGKICTVYGWNT-QYGGQAGVLEARVPIISNDV 321
Qy 354 CNADAYQGEVTEKMACAGIPREGVDTCCGDSGGPLMTQSDQMHVGVISWGYCGGPGSTPGVTVKVS 421
Db 322 CNGAFYGNQIKPKMFCAGYPEGGIDACGDSGGPFVCEDSISRTPRMRLCGIVSWGTGTC 381
Qy 409 GGPSTPGVTVKVSAYLNMTYVWK 432

Db 382 ALAQKFGVTVKVSDFREMIPOAIK 405

RESULT 3
US-11-182-752-2
; Sequence 2, Application US/11182752
; Publication No. US20050250154A1
; GENERAL INFORMATION:
; APPLICANT: Gan, Weiniu et al.
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: CL001194DIVII
; CURRENT APPLICATION NUMBER: US/11/182,752
; CURRENT FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 376
; TYPE: PRT
; ORGANISM: HUMAN
US-11-182-752-2

Query Match 24.8%; Score 580; DB 7; Length 376;
Best Local Similarity 30.8%; Pred. No. 4.8e-51;
Matches 132; Conservative 57; Mismatches 146; Indels 94; Gaps 9;

Qy 19 KPRIMEFRKVGIPITIALSLASITIVVLIKVIDKYFLCGQPLHFPKQQLDGE 78
Db 15 RPKVAALT---AGTLILLTAIGASMAIVAVLLR----- 45
Qy 79 LDCPLGDEDEHCVKSPFEGPAVAVRLSKDRSTLOVLDATGNWFSACFDNFTALAEATAC 138
Db 46 -----SDQE-----PLYPVQVSSADARLMVFDKTEGTWRLLCSSRSNARVAGLSC 90
Qy 139 ROMGYSSKPTFAVEIGPDOLDVETEN-----SOELRMKNSGPGCLS 183
Db 91 EEMGF-----LRAL---THSELDTVATAGNIGSFVFCVDEBGLPHTQHLLEVITSPDCPR 142
Qy 198 LKTPRVVGGEEASVDSWPMQVSIQYDKQHVCGSILDPHWLTAHCFRKHDTVDYFN-KKV 256
Db 117 LPVDRIVGGRDTSIGRPMQVSLRYDGAHLCGSLLSGDWVLTAAHCFRKHDTVDYFN-KKV 176
Qy 257 RAGSDKLSFPSLAVAKIIIEFNMPY-----KNDIALMKLOPFLTFSGTVAPIC 308
Db 177 FAGVAQASPHGLQVQAVVYHGGYLPFRDPNSENSNDIALVHLSPLRLEYIQPV 236
Qy 309 LPFDEBELTPATPLMIIGWFTKONGKMSDILLQASVQYIDSTRCNADDAAYQGEVTEK 368
Db 237 LPAAGQALVVGKICTVYGWNT-QYGGQAGVLEARVPIISNDVCGADFGNQIKPKM 295
Qy 369 MCAGIPREGVDTCCGDSGGPLMTQSDQMHVGVISWGYCGGPGSTPGVTVKVSAY 423
Db 296 FCAGYPEGGIDACGDSGGPFVCEDSISRTPRMRLCGIVSWGTGALAQKFGVTVKVSDF 355
Qy 424 LNMITYVWK 432
Db 356 REMIFQAIK 364

RESULT 4
US-11-067-811-4
; Sequence 4, Application US/11067811
; Publication No. US20050260688A1
; GENERAL INFORMATION:
; APPLICANT: Morgan, Bruce A
; APPLICANT: Enshell-SciJffers, David
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR HAIR GROWTH
; FILE REFERENCE: 10287-083001
; CURRENT APPLICATION NUMBER: US/11/067,811
; CURRENT FILING DATE: 2005-02-28
; PRIOR APPLICATION NUMBER: US 60/548,272
; PRIOR FILING DATE: 2004-02-27

QY	62	CGOPLHPIPRKOLCDGBELDCPLGEBDEBHCVKS	PEGPAAVAVRLSKDRSTLQVLDSATGNW	121
Db	64	CGOPLHPIPRKOLCDGBELDCPLGEBDEBHCVKS	PEGPAAVAVRLSKDRSTLQVLDSATGNW	123
QY	122	FSACFDNFTALAEACRQMGYSKPTFRAVEIGPDQDL	DVEITENSQELMRNNSGPGC	181
Db	124	FSACFDNFTALAEACRQMGYSKPTFRAVEIGPDQDL	DVEITENSQELMRNNSGPGC	183
QY	182	LSSGLVSLHCLACGSLKTPRVGGEASVDSWPWQVS	IOYDKOHVCGSILDPHVLTA	241
Db	184	LSSGLVSLHCLACGSLKTPRVGGEASVDSWPWQVS	IOYDKOHVCGSILDPHVLTA	243
QY	242	AHCFRKHTDVFNWKVRAGSDKLSFSLAVAKIIIEFNPMY	PKNDIALMKLOPPLTFS	301
Db	244	AHCFRKHTDVFNWKVRAGSDKLSFSLAVAKIIIEFNPMY	PKNDIALMKLOPPLTFS	303
QY	302	GTVRPICLPFDEBELTPATPLMIIGWFTKONGCKMSD	ILLOASVOYIDSTRCNADDAVQ	361
Db	304	GTVRPICLPFDEBELTPATPLMIIGWFTKONGCKMSD	ILLOASVOYIDSTRCNADDAVQ	363
QY	362	GEVTEKMMCAGIPREGVDTCQDGGSLMYSQDQWVIV	ISWGYCGGSPSTPGVYTKVS	421
Db	364	GEVTEKMMCAGIPREGVDTCQDGGSLMYSQDQWVIV	ISWGYCGGSPSTPGVYTKVS	423
QY	422	AYLNMIVYVWKAEL	435	
Db	424	AYLNMIVYVWKAEL	437	

Search completed: December 5, 2005, 13:49:12
 Job time : 192 secs

PT Detecting an ovarian cancer-associated transcript in a cell from a
 PT patient, comprises contacting a biological sample from the patient with a
 XX polynucleotide that hybridizes to an ovarian cancer gene.

PS Claim 13; Page 305; 332pp; English.

XX The invention relates to a method of detecting an ovarian cancer-
 CC associated transcript in a cell from a patient, by contacting a
 CC biological sample from the patient with a polynucleotide that selectively
 CC hybridizes to a sequence at least 80% identical to any of one of 80
 CC nucleic acid sequences given in the specification. The method is useful
 CC in diagnosing ovarian cancer and in identifying and using agents and/or
 CC targets that inhibit ovarian cancer. The nucleic acid molecule,
 CC polypeptide and the antibody may also be used in detecting ovarian
 CC cancer, monitoring and early detection of relapse following treatment,
 CC monitoring response to therapy, selecting patients for post-operative
 CC chemotherapy or radiation therapy, in selecting mode of therapy,
 CC determining tumour prognosis, early detection of pre-cancerous lesions,
 CC and as vaccines. This sequence corresponds to one of the proteins used
 CC for the detection method of the invention.

XX Sequence 437 AA;

Query Match 99.8%; Score 2337; DB 7; Length 437;
 Best Local Similarity 100.0%; Pred. No. 1.6e-201;

Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	2	DPDSQPLNSLDVPLRKPRIPMETFRKVGIPITIALSLASIIIVVLKVIIDKXYFL	61
DB	4	DPDSQPLNSLDVPLRKPRIPMETFRKVGIPITIALSLASIIIVVLKVIIDKXYFL	63
QY	62	CGQPLHIFPRKQLCDGELDCPLGDEDEHCVKSPFEGSAVAVRLSKDRSTLQVLDATGNW	121
DB	64	CGQPLHIFPRKQLCDGELDCPLGDEDEHCVKSPFEGSAVAVRLSKDRSTLQVLDATGNW	123
QY	122	FSACFDNFTALACTACROWGSSKPFPAVEIGPDDDLVETITNSQELPRNNSGRC	181
DB	124	FSACFDNFTALACTACROWGSSKPFPAVEIGPDDDLVETITNSQELPRNNSGRC	183
QY	182	LSGSLVSLHCLACGKSLKTPRVVVGEEASVDSWPMQVSIQYDKQHCGSILDPHVLTA	241
DB	184	LSGSLVSLHCLACGKSLKTPRVVVGEEASVDSWPMQVSIQYDKQHCGSILDPHVLTA	243
QY	242	AHCFRKHTDVFNKVRASGDKLSPSLNAKIIIEFNPMYKCDIALMLQLPFLFS	301
DB	244	AHCFRKHTDVFNKVRASGDKLSPSLNAKIIIEFNPMYKCDIALMLQLPFLFS	303
QY	302	GTVRPICLPFDEELTPATPLWIIIGWFTKONGKMSDILQASVQVIDSTRCNADDAVQ	361
DB	304	GTVRPICLPFDEELTPATPLWIIIGWFTKONGKMSDILQASVQVIDSTRCNADDAVQ	363
QY	362	GEVTEKMGACGPGEGVDTCCGDSGGPLMYQSDQMHVGVISWGYGCGGPTPGVYTKVS	421
DB	364	GEVTEKMGACGPGEGVDTCCGDSGGPLMYQSDQMHVGVISWGYGCGGPTPGVYTKVS	423
QY	422	AYLNMWYVWKAEL	435
DB	424	AYLNMWYVWKAEL	437

RESULT 15

ADN39878 standard; protein; 437 AA.

ADN39878;

17-UTN-2004 (first entry)

Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO: C248.

Human; differential expression; cancer; angiogenic disorder;

KM fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
 KM inflammatory disease; autoimmune disease;

KM retinal neovascularisation syndrome; scarring; uterine fibroid;
 KM detection; diagnosis; prognosis; drug screening; drug targeting;
 KM wound healing; contraception; cytostatic; cardiant; immunomodulatory;
 KM vulnery; gene therapy; vaccine.

OS Homo sapiens.

XX NO2003042661-A2.

XX 22-MAY-2003.

XX 13-NOV-2002; 2002MO-US036810.

XX 13-NOV-2001; 2001US-0350666P.

XX 29-NOV-2001; 2001US-0332464P.

XX 03-DEC-2001; 2001US-033394P.

XX 14-DEC-2001; 2001US-0340376P.

XX 08-JAN-2002; 2002US-0347211P.

XX 10-JAN-2002; 2002US-0347349P.

XX 08-FEB-2002; 2002US-0355250P.

XX 13-FEB-2002; 2002US-0356714P.

XX 20-FEB-2002; 2002US-0359077P.

XX 29-MAR-2002; 2002US-0368099P.

XX 04-APR-2002; 2002US-0370110P.

XX 12-APR-2002; 2002US-0372246P.

XX 05-JUN-2002; 2002US-0386614P.

XX 16-JUL-2002; 2002US-0396839P.

XX 22-JUL-2002; 2002US-0397775P.

XX 22-JUL-2002; 2002US-0397845P.

XX 09-SEP-2002; 2002US-0409450P.

XX (EOSB-) EOS BIOTECHNOLOGY INC.

XX Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevez PA;

XX Mack DH, Murray R, Watson SR, Wilson KE, Zlotnick A;

XX MPI, 2003-468649/44.

XX N-PSDB; ADN39861.

XX Determining the presence or absence of a pathological cell in a patient,

XX useful for diagnosing, prognosing or treating cancer, comprises detecting

XX a nucleic acid in a biological sample.

XX Claim 12; SEQ ID NO C248; 1385pp; English.

XX The invention relates to nucleic acids and proteins (ADN398683-ADN40064)

XX whose expression is upregulated or downregulated in specific cancers or

XX other diseases such as angiogenic or fibrotic disorders, and to methods

XX of determining the presence or absence of a pathological cell in a

XX patient by detecting a nucleic acid at least 80% identical to those of

XX the invention or by detecting a polypeptide of the invention. The

XX nucleic acid of the invention; antibodies which specifically bind a

XX polypeptide of the invention; use of such antibodies for drug targeting;

XX and methods of screening for modulators of activity or expression of the

XX polypeptides and nucleic acids. The nucleic acids, polypeptides,

XX antibodies and methods are useful for diagnosing, prognosing and treating

XX cancer and other conditions such as psoriasis, ischaemia, heart disease,

XX atherosclerosis, inflammatory diseases, autoimmune diseases, retinal

XX neovascularisation syndromes, scarring and uterine fibroids. They may

XX also be useful in wound healing and in contraception. The present

XX sequence represents a polypeptide of the invention.

XX Sequence 437 AA;

Query Match 99.8%; Score 2337; DB 7; Length 437;
 Best Local Similarity 100.0%; Pred. No. 1.6e-201;

Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	2	DPDSQPLNSLDVPLRKPRIPMETFRKVGIPITIALSLASIIIVVLKVIIDKXYFL	61
DB	4	DPDSQPLNSLDVPLRKPRIPMETFRKVGIPITIALSLASIIIVVLKVIIDKXYFL	63

QY 362 GEVTEKMMKAGIPGSGVDTCCGDSGGPLMYOSDQMHVVGIVSMGVCGGSPSTPGVYTKVS 421
Db 364 GEVTEKMMKAGIPGSGVDTCCGDSGGPLMYOSDQMHVVGIVSMGVCGGSPSTPGVYTKVS 423
QY 422 AYLNMIYVWKAEL 435
Db 424 AYLNMIYVWKAEL 437

RESULT 13
ABU04915 standard; protein; 437 AA.
XX ABU04915;
AC
XX 29-JAN-2003 (first entry)
DT
XX Human expressed protein tag (EPT) #1581.
DE
XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
KW proteinase; protease inhibitor; transporter; cytoskeletal protein;
KW receptor; transcription factor; cancer; MHC;
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX
OS Homo sapiens.
PN MO200278524-A2.
XX
PD 10-OCT-2002.
XX
PF 28-MAR-2002; 2002WO-US009671.
XX
PR 28-MAR-2001; 2001US-0279495P.
PR 21-MAY-2001; 2001US-0292544P.
PR 08-AUG-2001; 2001US-0310801P.
PR 01-OCT-2001; 2001US-0328370P.
PR 04-DEC-2001; 2001US-0336780P.
PR 20-FEB-2002; 2002US-0358985P.
XX
PA (ZYCO-) ZYCOS INC.
PI Chicx RM, Tomlinson AJ, Urban RG;
XX
DR WPI; 2003-040607/03.
XX
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
PT cytoskeletal proteins, receptors or transcription factors), useful for
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
PT leukemia.
XX
XX Example 2; SEQ ID NO 1581; 134pp; English.
PS
XX The invention describes a purified polypeptide, which comprises a
CC fragment of a kinase, phosphatase, protease, protease inhibitor,
CC transporter, cytoskeletal protein, receptor or transcription factor. The
CC polypeptide is useful as an immunogenic composition for eliciting in a
CC mammal an immunogenic response directed against any of the purified
CC polypeptide. The purified polypeptide, or the antibody that binds to this
CC polypeptide, is useful for treating cancer. The polypeptide is also
CC useful for identifying compounds that binds to a naturally processed
CC class I or class II MHC-binding polypeptide. The polypeptides and
CC polynucleotides are particularly useful for treating or preventing
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
CC lymphoma or leukaemia. These are also useful for screening agents for
CC treating the above mentioned diseases. This sequence represents an
CC expressed protein tag (EPT) isolated from human tissue for translational
CC profiling. Note: This sequence does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 437 AA;

Query Match 99.8%; Score 2337; DB 6; Length 437;
Best Local Similarity 100.0%; Pred. No. 1.6e-201;
Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 DPDSQPLNSLDVPLKRPRIPEMTFRKVGIPITIIALLSLASIIIVVLKIVLIDKTYFL 61
Db 4 DPDSQPLNSLDVPLKRPRIPEMTFRKVGIPITIIALLSLASIIIVVLKIVLIDKTYFL 63
QY 62 CGQPLHFIPIPRKQLCDGLDPLGDEDEHCYKSPFEGPAVAVRLSKDSTQVLDASGNW 121
Db 64 CGQPLHFIPIPRKQLCDGLDPLGDEDEHCYKSPFEGPAVAVRLSKDSTQVLDASGNW 123
QY 122 FSACFDNFTETALATACRQMGYSKPTFRAVEIGPDQDLVETITENSQELRMNSSGPC 181
Db 124 FSACFDNFTETALATACRQMGYSKPTFRAVEIGPDQDLVETITENSQELRMNSSGPC 183
QY 182 LSGSLVSLHCLACGSKLTPRVVGGEBASVDSWPQVSIQYDKQHVCGGSLDPHWVLT 241
Db 184 LSGSLVSLHCLACGSKLTPRVVGGEBASVDSWPQVSIQYDKQHVCGGSLDPHWVLT 243
QY 242 AHCRKHTDVFNWKVRAGSDKLSFSLAVAKIIIIIEPNMYPKONDIALMKLOPILTF 301
Db 244 AHCRKHTDVFNWKVRAGSDKLSFSLAVAKIIIIIEPNMYPKONDIALMKLOPILTF 303
QY 302 GTVRPILCPPEDELTATPLMTIIGWGTFTKONGKMSDIILOASVOYIDTRCNADAYQ 361
Db 304 GTVRPILCPPEDELTATPLMTIIGWGTFTKONGKMSDIILOASVOYIDTRCNADAYQ 363
QY 362 GEVTEKMMKAGIPGSGVDTCCGDSGGPLMYOSDQMHVVGIVSMGVCGGSPSTPGVYTKVS 421
Db 364 GEVTEKMMKAGIPGSGVDTCCGDSGGPLMYOSDQMHVVGIVSMGVCGGSPSTPGVYTKVS 423
QY 422 AYLNMIYVWKAEL 435
Db 424 AYLNMIYVWKAEL 437

RESULT 14
ADB80525 standard; protein; 437 AA.
XX
XX ADB80525;
AC
XX 04-DEC-2003 (first entry)
DT
XX
DE Ovarian cancer-associated protein #45.
XX
KW cytostatic; gene therapy; vaccine; ovarian cancer; diagnosis;
KW post-operative chemotherapy; radiation therapy; tumour prognosis;
KW pre-cancerous lesion detection.
XX
OS Homo sapiens.
PN MO2002102235-A2.
XX
PD 27-DEC-2002.
XX
PF 18-JUN-2002; 2002WO-US019297.
XX
PR 18-JUN-2001; 2001US-0299234P.
PR 27-AUG-2001; 2001US-0315287P.
PR 05-SEP-2001; 2001US-0317544P.
PR 13-NOV-2001; 2001US-0350666P.
PR 12-APR-2002; 2002US-0372246P.
XX
PA (BOSB-) BOS BIOTECHNOLOGY INC.
XX
XX Mack DH, Gish KC;
PI
XX WPI; 2003-167431/16.
DR
XX N-PSDB; ADB80524.
XX

transporter, cytoskeletal protein, receptor or transcription factor. The polypeptide is useful as an immunogenetic composition for eliciting in a mammal an immunogenic response directed against any of the purified polypeptide. The purified polypeptide, or the antibody that binds to this polypeptide, is useful for treating cancer. The polypeptide is also useful for identifying compounds that binds to a naturally processed class I or class II MHC-binding polypeptide. The polypeptides and polynucleotides are particularly useful for treating or preventing myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma, lymphoma or leukaemia. These are also useful for screening agents for treating the above mentioned diseases. This sequence represents an expressed protein tag (EPT) isolated from human tissue for translational profiling. Note: This sequence does not appear in the printed specification but was obtained in electronic format directly from <http://www.ncbi.nlm.nih.gov/pubmed/16444444>

Query Match	99.8%	Score 2337	DB 6	Length 437
Best Local Similarity	100.0%	Pred. No. 1.0e-201		
Matches 434	Conservative	0	Mismatches	0
			Indels	0
			Gaps	0

QY	DPDSQOPINSLDVYKLEKKPRIPMETPFKKVGPIIILALSLASIIIVVLIKVILDKYYFL	61
Db	4 DPDSQOPINSLDVYKLEKKPRIPMETPFKKVGPIIILALSLASIIIVVLIKVILDKYYFL	63
QY	62 CGOPLHPIPRKQICDGEIDCPLGEBDEBCHYKSPFEGPAVAVRLSKDRSTLOVLSATGMW	122
Db	64 CGOPLHPIPRKQICDGEIDCPLGEBDEBCHYKSPFEGPAVAVRLSKDRSTLOVLSATGMW	123
QY	122 FSACFDNFTMLATACRQMGYSKKPFRAVEIGPDODLVEIITENSQELMRNSSGPC	181
Db	124 FSACFDNFTMLATACRQMGYSKKPFRAVEIGPDODLVEIITENSQELMRNSSGPC	183
QY	182 LSGSLVSLHCLACGCKSLKTPRVVVGEBEASVDSMPWVSIOYKHVCGGSLIDDPHWLTA	241
Db	184 LSGSLVSLHCLACGCKSLKTPRVVVGEBEASVDSMPWVSIOYKHVCGGSLIDDPHWLTA	243
QY	242 AHCRKHTDVFNMKVRAGSDKLGSPSLAVAKIIIEFNPMPYKXNDIALMKLOPPLTFS	301
Db	244 AHCRKHTDVFNMKVRAGSDKLGSPSLAVAKIIIEFNPMPYKXNDIALMKLOPPLTFS	303
QY	302 GTVAPICLPFDEBLTPATPLMIIGMGFTKQNGKXSDILLQASVOVITDSTRCANADAYQ	361
Db	304 GTVAPICLPFDEBLTPATPLMIIGMGFTKQNGKXSDILLQASVOVITDSTRCANADAYQ	363
QY	362 GEVEYKMKMGIPFEGGVDTCCQSGSGPPLMYQSDQHWVGIYSWGVCGGSGPSTPGVYTKVS	421
Db	364 GEVEYKMKMGIPFEGGVDTCCQSGSGPPLMYQSDQHWVGIYSWGVCGGSGPSTPGVYTKVS	423
QY	422 AYLNMIYVWKAEL 435	
Db	424 AYLNMIYVWKAEL 437	

RESULT 12
ABU04936
ID ABU04936 standard; protein; 437 AA.

OS Homo sapiens.
XX

PN	WO200278524-A2.
XX	
PD	10-OCT-2002.
XX	
PF	28-MAR-2002; 2002WO-US009671.
XX	
PR	28-MAR-2001; 2001US-0279495P.
PR	21-MAY-2001; 2001US-0292544P.
PR	08-AUG-2001; 2001US-0310801P.
PR	01-OCT-2001; 2001US-0326370P.
PR	04-DEC-2001; 2001US-0336798P.
PR	20-FEB-2002; 2002US-0358985P.
XX	
XX	
PA	(ZYCO-) ZYCOS INC.
XX	
PI	Chica RM, Tomlinson AJ, Urban RG;
XX	
DR	WPI; 2003-040607/03.
XX	
PT	New polypeptides (e.g. kinases, phosphatases, proteases, transporters, cytoskeletal proteins, receptors or transcription factors), useful for treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or leukemia.
PT	

Query Match	99.8%	Score 2337	DB 6	Length 437
Best Local Similarity	100.0%	Pred. No. 1,6e-201		
Matches 434; Conservative	0	Mismatches	0	Gaps 0

QY	2	DPDSDGCLNSLDVKKPLRKRIIPMETPRKXGPIIIIALSLASIIIVVVLIVLIDKXYFL	61
Db	4	DPDSDGCLNSLDVKKPLRKRIIPMETPRKXGPIIIIALSLASIIIVVVLIVLIDKXYFL	63
QY	62	CGOPLHPIPKOJCDJELDCLPGDEBHCVKSFPBGPAVAVRLSKDRSTLQVLSATGNW	1221
Db	64	CGOPLHPIPKOJCDJELDCLPGDEBHCVKSFPBGPAVAVRLSKDRSTLQVLSATGNW	1223
QY	122	FSACFDFNFALAEFTACROMGYSKXPTPAVEIGPDOLDVYEITENSQELRMNRSSGPC	1871
Db	124	FSACFDFNFALAEFTACROMGYSKXPTPAVEIGPDOLDVYEITENSQELRMNRSSGPC	1873
QY	182	LGSGLVSLHCLACGSKLTPRVVGBEASVDSMPQVSIQYDKQHCVCSSILDDHWWLTA	2411
Db	184	LGSGLVSLHCLACGSKLTPRVVGBEASVDSMPQVSIQYDKQHCVCSSILDDHWWLTA	2413
QY	242	AHCERKHTDVENMKKAPAGSDKLGSPSLAVAKIIIEENPMYPKXNDIATMLKLOPLTFPS	3011
Db	244	AHCERKHTDVENMKKAPAGSDKLGSPSLAVAKIIIEENPMYPKXNDIATMLKLOPLTFPS	3013
QY	302	GIVRPICLDFPDEBELPATPLWIIQWGEFTKQNGKMSDIIILQASVOVLDSTRCANADAYO	3611
Db	304	GIVRPICLDFPDEBELPATPLWIIQWGEFTKQNGKMSDIIILQASVOVLDSTRCANADAYO	3613

ABU04920
ID ABU04920 standard; protein; 437 AA.
AC ABU04920;
XX
DT 29-JAN-2003 (first entry)
XX
DE Human expressed protein tag (EPT) #1586.
XX
KM Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
KM protease; protease inhibitor; transporter; cytoskeletal protein;
KM receptor; transcription factor; cancer; MHC;
KM major histocompatibility complex; myeloma; colon cancer; gastric cancer;
KM adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX
OS Homo sapiens.
FN WO200278524-A2.
PD 10-OCT-2002.
XX
PE 28-MAR-2002; 2002WO-US009671.
XX
PR 28-MAR-2001; 2001US-0279495P.
PR 21-MAY-2001; 2001US-0292544P.
PR 08-AUG-2001; 2001US-0310801P.
PR 01-OCT-2001; 2001US-0326370P.
PR 04-DEC-2001; 2001US-0336780P.
PR 20-FEB-2002; 2002US-0358985P.
XX
PA (ZYCO-) ZYCOS INC.
PI Chicx RM, Tomlinson AJ, Urban RG;
XX
DR WPI; 2003-040607/03.
XX
PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
PT cytoskeletal proteins, receptors or transcription factors), useful for
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
PT leukemia.
XX
PS Example 2; SEQ ID NO 1586; 134pp; English.
XX
CC The invention describes a purified polypeptide, which comprises a
CC fragment of a kinase, phosphatase, protease, protease inhibitor
CC transporter, cytoskeletal protein, receptor or transcription factor. The
CC polypeptide is useful as an immunogenic composition for eliciting in a
CC mammal an immunogenic response directed against any of the purified
CC polypeptide. The purified polypeptide, or the antibody that binds to this
CC polypeptide, is useful for treating cancer. The polypeptide is also
CC useful for identifying compounds that binds to a naturally processed
CC class I or class II MHC-binding polypeptide. The polypeptides and
CC polynucleotides are particularly useful for treating or preventing
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
CC lymphoma or leukaemia. These are also useful for screening agents for
CC treating the above mentioned diseases. This sequence represents an
CC expressed protein tag (EPT) isolated from human tissue for translational
CC profiling. Note: This sequence does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 437 AA;

Query Match 99.8%; Score 2337; DB 6; Length 437;
Best Local Similarity 100.0%; Pred. No. 1.6e-201;
Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DPDSQPIINSLDVKRLKRPIMETPRKVGIPPIIALSLASIIIVVILKYLKYYFL 61
DB 4 DPDSQPIINSLDVKRLKRPIMETPRKVGIPPIIALSLASIIIVVILKYLKYYFL 63
QY 62 CGOPLHFI PRKQLCDGELDCPLGDEDEHCVKSPFPGPAVAVALSKDRSTLQVLD SATGNW 121

DB 64 CGOPLHFI PRKQLCDGELDCPLGDEDEHCVKSPFPGPAVAVALSKDRSTLQVLD SATGNW 123
QY 122 FSACFDNFTALAEACRQNGYSKPTFAVEIGPDODLVEITENSQELRMNSSGPC 181
DB 124 FSACFDNFTALAEACRQNGYSKPTFAVEIGPDODLVEITENSQELRMNSSGPC 183
QY 182 LSGSLVSLHCLACGSKSLKTPRVVGGEEASVDSMFQVSIQYDKQHVCGSILDPHWLTA 241
DB 184 LSGSLVSLHCLACGSKSLKTPRVVGGEEASVDSMFQVSIQYDKQHVCGSILDPHWLTA 243
QY 242 AHCFRKTDPVNMKVRGSGSLGSPSLAVAKIIIEFNMPYKPDNIIALMKLOEPLTFS 301
DB 244 AHCFRKTDPVNMKVRGSGSLGSPSLAVAKIIIEFNMPYKPDNIIALMKLOEPLTFS 303
QY 302 GTVAPICLPFDEELTPATPLMIIGWFTKONGGKMSDILQASVQYIDSTRCNADDAAYQ 361
DB 304 GTVAPICLPFDEELTPATPLMIIGWFTKONGGKMSDILQASVQYIDSTRCNADDAAYQ 363
QY 362 GEVTERKMCAGIPRGVDTCCGDSGFLMYOSDQMHVGVISWVGCGGSPSTPCVYTKVS 421
DB 364 GEVTERKMCAGIPRGVDTCCGDSGFLMYOSDQMHVGVISWVGCGGSPSTPCVYTKVS 423
QY 422 AYLMWITVWKAEL 435
DB 424 AYLMWITVWKAEL 437
RESULT 11
ABU04935
ID ABU04935 standard; protein; 437 AA.
XX
AC ABU04935;
XX
DT 29-JAN-2003 (first entry)
XX
DE Human expressed protein tag (EPT) #1601.
XX
KM Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
KM protease; protease inhibitor; transporter; cytoskeletal protein;
KM receptor; transcription factor; cancer; MHC;
KM major histocompatibility complex; myeloma; colon cancer; gastric cancer;
KM adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX
OS Homo sapiens.
FN WO200278524-A2.
PD 10-OCT-2002.
XX
PE 28-MAR-2002; 2002WO-US009671.
XX
PR 28-MAR-2001; 2001US-0279495P.
PR 21-MAY-2001; 2001US-0292544P.
PR 08-AUG-2001; 2001US-0310801P.
PR 01-OCT-2001; 2001US-0326370P.
PR 04-DEC-2001; 2001US-0336780P.
PR 20-FEB-2002; 2002US-0358985P.
XX
PA (ZYCO-) ZYCOS INC.
PI Chicx RM, Tomlinson AJ, Urban RG;
XX
DR WPI; 2003-040607/03.
XX
PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
PT cytoskeletal proteins, receptors or transcription factors), useful for
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
PT leukemia.
XX
PS Example 2; SEQ ID NO 1601; 134pp; English.
XX
CC The invention describes a purified polypeptide, which comprises a
CC fragment of a kinase, phosphatase, protease, protease inhibitor.

XX	Sequence 435 AA;	
SO		
Query Match	99.8%; Score 2338; DB 6; Length 435;	
Best Local Similarity	99.8%; Pred. No. 1.3e-201;	
Matches 434; Conservative	1; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MDPSDQPLNSLDVYPLKPRIPMETFRKVGIPPIIALSLASIIIVVLIKVILDKYF 60	
DB	1 MDPSDQPLNSLDVYPLKPRIPMETFRKVGIPPIIALSLASIIIVVLIKVILDKYF 60	
QY	61 LCQGPLHPIPRKQICDGLDCEPLGDEDEHCVKSPFEGPAVAVRLSKRSTLQVLDATGN 120	
DB	61 LCQGPLHPIPRKQICDGLDCEPLGDEDEHCVKSPFEGPAVAVRLSKRSTLQVLDATGN 120	
QY	121 WFSACFDNFTALAEACRQWYSSKPTFRAVEIGPQDDLDVVEITENSOELMRNSSGP 180	
DB	121 WFSACFDNFTALAEACRQWYSSKPTFRAVEIGPQDDLDVVEITENSOELMRNSSGP 180	
QY	181 CLSGSLVSLHCLACGSKLTPRVVGGEBASVDSMPQVSIQYDKQHVCGGSILDPHWLVT 240	
DB	181 CLSGSLVSLHCLACGSKLTPRVVGGEBASVDSMPQVSIQYDKQHVCGGSILDPHWLVT 240	
QY	241 AAHCFRKHITVFNWKVAVAGSDKLGSPSLAVAKIIIEFNPMYPKNDIALMLKQPLTF 300	
DB	241 AAHCFRKHITVFNWKVAVAGSDKLGSPSLAVAKIIIEFNPMYPKNDIALMLKQPLTF 300	
QY	301 SGTVRPICLPFPDEBELTPATPLMIIGWFTKQNGKMSDILLQASVQVIDSTRCNADAY 360	
DB	301 SGTVRPICLPFPDEBELTPATPLMIIGWFTKQNGKMSDILLQASVQVIDSTRCNADAY 360	
QY	361 QGEVTEKMMGAGIPEGGVDTCCGDSGGPLMYQSDQHWVGIYSWGYCGGSPSTPGVYTKY 420	
DB	361 QGEVTEKMMGAGIPEGGVDTCCGDSGGPLMYQSDQHWVGIYSWGYCGGSPSTPGVYTKY 420	
QY	421 SAYLNMWYVWKAEI 435	
DB	421 SAYLNMWYVWKAEI 435	
RESULT 9		
AAG79359	standard; protein; 437 AA.	
ID	AAG79359	
AC	AAG79359;	
XX		
DT	21-AUG-2002 (first entry)	
XX		
DE	CJ48 preferred sequence.	
XX		
KW	Colorectal cancer; CGA7; CJ48; modulating protein; screening;	
XX		
OS	drug candidate; vaccine.	
XX		
OS	Homo sapiens.	
XX		
FT	Key	Location/Qualifiers
FT	Domain	36..63
FT		/note= "Transmembrane domain"
XX		
PN	US2002042067-A1.	
XX		
PD	11-APR-2002.	
XX		
PF	08-MAY-2001; 2001US-00851586.	
XX		
PR	17-AUG-2000; 2000US-00642252.	
XX		
PR	06-SEP-2000; 2000US-00656002.	
XX		
PA	(MACK/) MACK D.	
PA	(GISH/) GISH K C.	
XX	(WILS/) WILSON K E.	
XX		
PI	Mack D, Gish KC, Wilson KE;	

XX	Sequence 437 AA;	
SO		
Query Match	99.8%; Score 2337; DB 5; Length 437;	
Best Local Similarity	100.0%; Pred. No. 1.6e-201;	
Matches 434; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
QY	2 DPDSDDQPLNSLDVYPLKPRIPMETFRKVGIPPIIALSLASIIIVVLIKVILDKYF 61	
DB	4 DPDSDDQPLNSLDVYPLKPRIPMETFRKVGIPPIIALSLASIIIVVLIKVILDKYF 63	
QY	62 CGQPLHPIPRKQICDGLDCEPLGDEDEHCVKSPFEGPAVAVRLSKRSTLQVLDATGN 121	
DB	64 CGQPLHPIPRKQICDGLDCEPLGDEDEHCVKSPFEGPAVAVRLSKRSTLQVLDATGN 123	
QY	122 FSACFDNFTALAEACRQWYSSKPTFRAVEIGPQDDLDVVEITENSOELMRNSSGP 181	
DB	124 FSACFDNFTALAEACRQWYSSKPTFRAVEIGPQDDLDVVEITENSOELMRNSSGP 183	
QY	182 LSGSLVSLHCLACGSKLTPRVVGGEBASVDSMPQVSIQYDKQHVCGGSILDPHWLVT 241	
DB	184 LSGSLVSLHCLACGSKLTPRVVGGEBASVDSMPQVSIQYDKQHVCGGSILDPHWLVT 243	
QY	242 AHCFRKHITVFNWKVAVAGSDKLGSPSLAVAKIIIEFNPMYPKNDIALMLKQPLTF 301	
DB	244 AHCFRKHITVFNWKVAVAGSDKLGSPSLAVAKIIIEFNPMYPKNDIALMLKQPLTF 303	
QY	302 GTVVPICLPFPDEBELTPATPLMIIGWFTKQNGKMSDILLQASVQVIDSTRCNADAY 361	
DB	304 GTVVPICLPFPDEBELTPATPLMIIGWFTKQNGKMSDILLQASVQVIDSTRCNADAY 363	
QY	362 GEVTEKMMGAGIPEGGVDTCCGDSGGPLMYQSDQHWVGIYSWGYCGGSPSTPGVYTKY 421	
DB	364 GEVTEKMMGAGIPEGGVDTCCGDSGGPLMYQSDQHWVGIYSWGYCGGSPSTPGVYTKY 423	
QY	422 AYLNWYVWKAEI 435	
DB	424 AYLNWYVWKAEI 437	
RESULT 10		

Db	301	SGTVAPICLPFEDDELTPTATPLMIIGWFTKONGKMSDILLQASVQVIDSTRCNADDA	360
QY	361	OGEVTEKMMKACGIPGSGVDTCGDSGGPLMYOSDQHWVGIVSWGVCGGSPSPGYTTK	420
Db	361	OGEVTEKMMKACGIPGSGVDTCGDSGGPLMYOSDQHWVGIVSWGVCGGSPSPGYTTK	420
QY	421	SAYLNMIVNWKAEI 435	
Db	421	SAYLNMIVNWKAEI 435	
RESULT 8			
ABU04930			
ID	ABU04930	standard; protein; 435 AA.	
XX			
AC	ABU04930;		
XX			
DT	29-JAN-2003	(first entry)	
XX			
DE	Human expressed protein tag (EPT) #1596.		
KM	Translational profiling; expressed protein tag; EPT; kinase; phosphatase;		
KM	protease; protease inhibitor; transporter; cytoskeletal protein;		
KM	receptor; transcription factor; cancer; MHC;		
KM	major histocompatibility complex; myeloma; colon cancer; gastric cancer;		
KM	adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.		
OS	Homo sapiens.		
XX			
PN	WO200278524-A2.		
XX			
PD	10-OCT-2002.		
XX			
EF	28-MAR-2002; 2002WO-US009671.		
XX			
PR	28-MAR-2001; 2001US-0279495P.		
PR	21-MAY-2001; 2001US-0292544P.		
PR	08-AUG-2001; 2001US-0310801P.		
PR	01-OCT-2001; 2001US-0326370P.		
PR	04-DEC-2001; 2001US-0336780P.		
PR	20-FEB-2002; 2002US-0358985P.		
XX			
PA	(ZYCO-) ZYCOS INC.		
XX			
PI	Chicz RM, Tomlinson AJ, Urban RG;		
XX			
DR	WPI; 2003-040607/03.		
XX			
PT	New polypeptides (e.g. kinases, phosphatases, proteases, transporters,		
PT	cytoskeletal proteins, receptors or transcription factors), useful for		
PT	treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or		
PT	leukemia.		
XX			
PS	Example 2; SEQ ID NO 1596; 134pp; English.		
XX			
CC	The invention describes a purified polypeptide, which comprises a		
CC	fragment of a kinase, phosphatase, protease, protease inhibitor,		
CC	transporter, cytoskeletal protein, receptor or transcription factor. The		
CC	polypeptide is useful as an immunogenic composition for eliciting in a		
CC	mammal an immunogenic response directed against any of the purified		
CC	polypeptide. The purified polypeptide, or the antibody that binds to this		
CC	polypeptide, is useful for treating cancer. The polypeptide is also		
CC	useful for identifying compounds that binds to a naturally processed		
CC	class I or class II MHC-binding polypeptide. The polypeptides and		
CC	polynucleotides are particularly useful for treating or preventing		
CC	myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,		
CC	lymphoma or leukemia. These are also useful for screening agents for		
CC	treating the above mentioned diseases. This sequence represents an		
CC	expressed protein tag (EPT) isolated from human tissue for translational		
CC	profiling. Note: This sequence does not appear in the printed		
CC	specification but was obtained in electronic format directly from WIPO at		
CC	ftp.wipo.int/pub/published pct sequences		

```

Db      301 SGTVPRLCLPFPDEBLTPATPLMIIGWFTKONGKMSDILLQASVQVIDSTRCNADDAAY 360
Qy      361 QGEVTEKMKACGIPESGVDTCCGDSGGPLMYGSDQMHVVGIVSWGCGGSPSTPGYTYK 420
Db      361 QGEVTEKMKACGIPESGVDTCCGDSGGPLMYGSDQMHVVGIVSWGCGGSPSTPGYTYK 420
Qy      421 SAYLWMIYVWKAEL 435
Db      421 SAYLWMIYVWKAEL 435

RESULT 6
AEA10749 standard; protein; 461 AA.
AEA10749;
AC      AEA10749;
DT      28-JUL-2005 (first entry)
DE      Mutated human Ovr115 protein #3.
XX
XX      Antibody therapy; Ovr115; ovarian tumor; breast tumor; neoplasm;
KM      cytosolic; muten; antibody.
XX
OS      Homo sapiens.
OS      Mus sp.
OS      Synthetic.
XX
XX      WO2005046573-A2.
PD      26-MAY-2005.
XX
XX      28-JUN-2004; 2004WO-US020741.
PF
XX
XX      27-JUN-2003; 2003US-0485346P.
PR      17-NOV-2003; 2003US-0523271P.
XX
XX      (DIAD-) DIADEXUS INC.
PA
XX
PI      Papheoff J, Pilkington G, Keller G, Li W, Corral L, Simon I;
PI      Kmet M, Tang U;
XX
XX      WPI; 2005-366911/37.
DR
XX
XX      New isolated Pro104 antibody that binds to Pro104 on a mammalian cell in
PT      vivo, useful for diagnosing and treating cancer particularly ovarian
PT      serous adenocarcinoma and breast infiltrating ductal carcinoma.
XX
XX      Example 1; SEQ ID NO 7; 213pp; English.
PS
XX
XX      The invention relates to an isolated Pro104 antibody that binds to Pro104
CC      on a mammalian cell in vivo. The invention also relates to a cell that
CC      produces the antibody, a method of producing the antibody, a composition
CC      comprising the antibody and a carrier, a method of killing a Pro104-
CC      expressing cancer cell, a method of alleviating a Pro104-expressing
CC      cancer in a mammal, a method for determining if cells in a sample express
CC      Pro104, a method for detecting Pro104 over expression in a test cell
CC      sample, a method for detecting Pro104 over-expression in a subject and a
CC      method of screening for antibodies that bind to an epitope which is bound
CC      by the antibody. The anti-Pro104 antibody is useful for diagnosing and
CC      treating Pro104-expressing cancers, particularly ovarian cancer and
CC      breast cancer. This sequence represents a mutated human Ovr115 protein,
CC      used in the scope of the invention.
XX
XX      Sequence 461 AA;
SQ
Query Match      100.0%; Score 2342; DB 9; Length 461;
Best local Similarity 100.0%; Pred. No. 6.2e-202;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1 MDPSDQPLNSLDVPLKRPRIEMETFRKVGIPITIALISLASIIIVVLIKVILDKYF 60

```

```

Db      1 MDPSDQPLNSLDVPLKRPRIEMETFRKVGIPITIALISLASIIIVVLIKVILDKYF 60
Qy      61 LCGQPLHFI PRKQLCDGBLDCPLGBDEBHCYKSPFEGPAVAVRLSKRSTLQVLSATGN 120
Db      61 LCGQPLHFI PRKQLCDGBLDCPLGBDEBHCYKSPFEGPAVAVRLSKRSTLQVLSATGN 120
Qy      121 WFSACFDNTEALATACRQMGYSKPTFRVAIEIGPDQLDVEITENSQELMRNMSGP 180
Db      121 WFSACFDNTEALATACRQMGYSKPTFRVAIEIGPDQLDVEITENSQELMRNMSGP 180
Qy      181 CLSGSLVSLHCLACGSLKTPRVGGEASVDSWPQVSIQYDKQHVCGGSLIDPMMVLT 240
Db      181 CLSGSLVSLHCLACGSLKTPRVGGEASVDSWPQVSIQYDKQHVCGGSLIDPMMVLT 240
Qy      241 AAHCRKHTDVNMKVRAGSDKLSFSPSLAVAKIIIEBNPMYPKNDIATMKLQPLTF 300
Db      241 AAHCRKHTDVNMKVRAGSDKLSFSPSLAVAKIIIEBNPMYPKNDIATMKLQPLTF 300
Qy      301 SGTVPRLCLPFPDEBLTPATPLMIIGWFTKONGKMSDILLQASVQVIDSTRCNADDAAY 360
Db      301 SGTVPRLCLPFPDEBLTPATPLMIIGWFTKONGKMSDILLQASVQVIDSTRCNADDAAY 360
Qy      361 QGEVTEKMKACGIPESGVDTCCGDSGGPLMYGSDQMHVVGIVSWGCGGSPSTPGYTYK 420
Db      361 QGEVTEKMKACGIPESGVDTCCGDSGGPLMYGSDQMHVVGIVSWGCGGSPSTPGYTYK 420
Qy      421 SAYLWMIYVWKAEL 435
Db      421 SAYLWMIYVWKAEL 435

RESULT 7
AA06437
ID      AA06437 standard; protein; 435 AA.
XX
XX      AA06437;
AC
XX
XX      27-SEP-1999 (first entry)
DT
XX
XX      Human protease HUPW-6.
DE
XX
XX      Serine protease; human; HUPW-6; cell proliferation; cancer;
KM      immune disorder; inflammation; therapy.
XX
XX      Homo sapiens.
OS
XX
XX      Key
FH      Key
FH      Location/Qualifiers
FT      Modified-site
FT      /note= "protein kinase C phosphorylation site"
FT      26
FT      Modified-site
FT      /note= "tyrosine kinase phosphorylation site"
FT      59
FT      Modified-site
FT      /note= "casein kinase II phosphorylation site"
FT      93
FT      Modified-site
FT      /note= "N-glycosylated"
FT      128
FT      Modified-site
FT      /note= "protein kinase C phosphorylation site"
FT      144
FT      Modified-site
FT      /note= "protein kinase C phosphorylation site"
FT      148
FT      Modified-site
FT      /note= "protein kinase C phosphorylation site"
FT      176
FT      Modified-site
FT      /note= "N-glycosylated"
FT      197
FT      Modified-site
FT      /note= "protein kinase C phosphorylation site"
FT      200
FT      Modified-site
FT      /note= "protein kinase C phosphorylation site"
FT      231
FT      Modified-site
FT      /note= "casein kinase II phosphorylation site"
FT      243
FT      Active-site
FT      /note= "characteristic of serine protease"
FT      249
FT      Modified-site
FT      /note= "cAMP- and cGMP-dependent protein kinase
FT      phosphorylation site"
FT      260

```

PR 01-OCT-2001; 2001US-0326370P.
 PR 04-DEC-2001; 2001US-0336780P.
 PR 20-FEB-2002; 2002US-0358985P.
 XX
 XX (ZYCO-) ZYCOs INC.
 PA
 XX Chicx RM, Tomlinson AJ, Urban RG;
 PI WPI; 2003-040607/03.
 XX
 XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
 PT cytoskeletal proteins, receptors or transcription factors), useful for
 PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
 PT leukemia.
 XX
 XX Example 2; SEQ ID NO 1578; 134pp; English.

CC The invention describes a purified polypeptide, which comprises a
 CC fragment of a kinase, phosphatase, protease, protease inhibitor,
 CC transporter, cytoskeletal protein, receptor or transcription factor. The
 CC polypeptide is useful as an immunogenic composition for eliciting in a
 CC mammal an immunogenic response directed against any of the purified
 CC polypeptide, the purified polypeptide, or the antibody that binds to this
 CC polypeptide, is useful for treating cancer. The polypeptide is also
 CC useful for identifying compounds that binds to a naturally processed
 CC class I or class II MHC-binding polypeptide. The polypeptides and
 CC polynucleotides are particularly useful for treating or preventing
 CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
 CC lymphoma or leukemia. These are also useful for screening agents for
 CC treating the above mentioned diseases. This sequence represents an
 CC expressed protein tag (EPT) isolated from human tissue for translational
 CC profiling. Note: This sequence does not appear in the printed
 CC specification but was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 435 AA;

Query Match 100.0%; Score 2342; DB 6; Length 435;
 Best Local Similarity 100.0%; Pred. No. 5.8e-202;
 Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPDSQPLNSLDVYKPKRPIPMETFRKVGIPITIALSLASIIIVVLKIVLDKYYF 60
 DB 1 MDPDSQPLNSLDVYKPKRPIPMETFRKVGIPITIALSLASIIIVVLKIVLDKYYF 60
 QY 61 LCGQPLHPIPRKQICDELCPDGEDEBHCVSFPEGPAVAVRISKDSTQVDSATGN 120
 DB 61 LCGQPLHPIPRKQICDELCPDGEDEBHCVSFPEGPAVAVRISKDSTQVDSATGN 120
 QY 121 WFSACFDFNFTALAEACRQWYSSKPTFRAVEIGPPDODLDVVEITENSQELRMNNSGP 180
 DB 121 WFSACFDFNFTALAEACRQWYSSKPTFRAVEIGPPDODLDVVEITENSQELRMNNSGP 180
 QY 181 CLSGSLVSLHCLACGSKSLKTPRVVGGEBASVDSWPMQVSIQYDKQHVCGGSIIDPHWLT 240
 DB 181 CLSGSLVSLHCLACGSKSLKTPRVVGGEBASVDSWPMQVSIQYDKQHVCGGSIIDPHWLT 240
 QY 241 AAHCFRKTIDVFNKVRAGSKLGSFPSLAIVAKIIIEFNPMYRKNDIAMLKQFPLTF 300
 DB 241 AAHCFRKTIDVFNKVRAGSKLGSFPSLAIVAKIIIEFNPMYRKNDIAMLKQFPLTF 300
 QY 301 SGTVRPILCPFDEELTPATPLMIIGWGFTKONGKMSDILLQASVOYIDSTRCNADAY 360
 DB 301 SGTVRPILCPFDEELTPATPLMIIGWGFTKONGKMSDILLQASVOYIDSTRCNADAY 360
 QY 361 QGEVTEKMCAGIPBEGVDTCQSDSGPPLMYQSDOMHWVGIIVSWGSGCGSPGVYTKV 420
 DB 361 QGEVTEKMCAGIPBEGVDTCQSDSGPPLMYQSDOMHWVGIIVSWGSGCGSPGVYTKV 420
 QY 421 SAYLNMIVNWKAEL 435
 DB 421 SAYLNMIVNWKAEL 435

RESULT 5
 ADU91838
 ID ADU91838 standard; protein; 461 AA.
 XX
 XX AC ADU91838;
 XX
 XX 24-FEB-2005 (first entry)
 DT
 XX
 XX Ovr115.
 DE
 XX
 XX cytosstatic; Ovr115.
 KW
 XX
 XX Homo sapiens.
 OS
 OS Tobacco etch virus.
 OS Unidentified.
 XX
 XX MO2004104173-A2.
 XX
 XX
 XX .02-DEC-2004.
 PD
 XX
 XX 17-MAY-2004; 2004MO-US015258.
 PF
 XX
 XX 16-MAY-2003; 2003US-0471068P.
 PR
 PR 05-APR-2004; 2004US-0559730P.
 XX
 XX (DIAD-) DIADEXUS INC.
 PA
 XX
 XX Pilkington G, Keller G, Li W, Corral L, Simon I;
 PI WPI; 2004-834291/82.
 DR
 XX
 XX New isolated anti-ovarian, pancreatic, lung or breast cancer antigen
 PT (Ovr115) antibody that binds to Ovr115 on a mammalian cell, useful for
 PT diagnosing or treating ovarian, pancreatic and colon cancer.
 PT
 XX
 XX Example 1; SEQ ID NO 3; 169pp; English.

CC The invention relates to an isolated anti-ovarian, pancreatic, lung or
 CC breast cancer antigen (Ovr115) antibody that binds to Ovr115 on a
 CC mammalian cell in vivo. The methods and compositions of the present
 CC invention are useful for producing anti-Ovr115 antibody compositions for
 CC diagnosing or treating Ovr115-expressing ovarian, pancreatic and colon
 CC cancer cells. The present sequence represents the amino acid sequence of
 CC Ovr115.

XX Sequence 461 AA;

Query Match 100.0%; Score 2342; DB 8; Length 461;
 Best Local Similarity 100.0%; Pred. No. 6.2e-202;
 Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPDSQPLNSLDVYKPKRPIPMETFRKVGIPITIALSLASIIIVVLKIVLDKYYF 60
 DB 1 MDPDSQPLNSLDVYKPKRPIPMETFRKVGIPITIALSLASIIIVVLKIVLDKYYF 60
 QY 61 LCGQPLHPIPRKQICDELCPDGEDEBHCVSFPEGPAVAVRISKDSTQVDSATGN 120
 DB 61 LCGQPLHPIPRKQICDELCPDGEDEBHCVSFPEGPAVAVRISKDSTQVDSATGN 120
 QY 121 WFSACFDFNFTALAEACRQWYSSKPTFRAVEIGPPDODLDVVEITENSQELRMNNSGP 180
 DB 121 WFSACFDFNFTALAEACRQWYSSKPTFRAVEIGPPDODLDVVEITENSQELRMNNSGP 180
 QY 181 CLSGSLVSLHCLACGSKSLKTPRVVGGEBASVDSWPMQVSIQYDKQHVCGGSIIDPHWLT 240
 DB 181 CLSGSLVSLHCLACGSKSLKTPRVVGGEBASVDSWPMQVSIQYDKQHVCGGSIIDPHWLT 240
 QY 241 AAHCFRKTIDVFNKVRAGSKLGSFPSLAIVAKIIIEFNPMYRKNDIAMLKQFPLTF 300
 DB 241 AAHCFRKTIDVFNKVRAGSKLGSFPSLAIVAKIIIEFNPMYRKNDIAMLKQFPLTF 300
 QY 301 SGTVRPILCPFDEELTPATPLMIIGWGFTKONGKMSDILLQASVOYIDSTRCNADAY 360

```
QY 241 AAHCRKHTDVFNWKVRAGSDKLGSPPSIAVAKIIIEFNPMYRKNDIALMKLQPLTF 300
DB 241 AAHCRKHTDVFNWKVRAGSDKLGSPPSIAVAKIIIEFNPMYRKNDIALMKLQPLTF 300
QY 301 SGTVRPICPFPEDELTPTPLMIIGWFTKONGKMSDILLQASVQVIDSTRCANADAY 360
DB 301 SGTVRPICPFPEDELTPTPLMIIGWFTKONGKMSDILLQASVQVIDSTRCANADAY 360
QY 361 QGEVTERKMKACGIPREGVDTCQDSDGSPLMYQSDQMHVGVISWGGCGSPSTPGVYTKV 420
DB 361 QGEVTERKMKACGIPREGVDTCQDSDGSPLMYQSDQMHVGVISWGGCGSPSTPGVYTKV 420
QY 421 SAYLNMWYVWKAEL 435
DB 421 SAYLNMWYVWKAEL 435

RESULT 3
ABU04931
ID ABU04931 standard; protein; 435 AA.
XX
AC ABU04931;
XX
DT 29-JAN-2003 (first entry)
XX
DE Human expressed protein tag (EPT) #1597.
XX
KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
KW protease; protease inhibitor; transporter; cytoskeletal protein;
KW receptor; transcription factor; cancer; MHC;
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200278524-A2.
XX
PD 10-OCT-2002.
XX
PF 28-MAR-2002; 2002WO-US009671.
XX
PR 28-MAR-2001; 2001US-0279495P.
PR 21-MAY-2001; 2001US-0292544P.
PR 08-AUG-2001; 2001US-0310801P.
PR 01-OCT-2001; 2001US-0326370P.
PR 04-DEC-2001; 2001US-0336780P.
PR 20-FEB-2002; 2002US-0358985P.
XX
PA (ZYCO-) ZYCOs INC.
XX
PI Chicx RM, Tomlinson AJ, Urban RG;
XX
DR WPI; 2003-040607/03.
XX
PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
PT cytoskeletal proteins, receptors or transcription factors), useful for
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
PT leukemia.
XX
PS Example 2; SEQ ID NO 1597; 134pp; English.
XX
OS The invention describes a purified polypeptide, which comprises a
CC fragment of a kinase, phosphatase, protease, protease inhibitor,
CC transporter, cytoskeletal protein, receptor or transcription factor. The
CC polypeptide is useful as an immunogenic composition for eliciting in a
CC mammal an immunogenic response directed against any of the purified
CC polypeptide. The purified polypeptide, or the antibody that binds to this
CC polypeptide, is useful for treating cancer. The polypeptide is also
CC useful for identifying compounds that binds to a naturally processed
CC class I or class II MHC-binding polypeptide. The polypeptides and
CC polynucleotides are particularly useful for treating or preventing
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
```

```
CC lymphoma or leukaemia. These are also useful for screening agents for
CC treating the above mentioned diseases. This sequence represents an
CC expressed protein tag (EPT) isolated from human tissue for translational
CC profiling. Note: This sequence does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 435 AA;
Query Match 100.0%; Score 2342; DB 6; Length 435;
Best Local Similarity 100.0%; Pred. No. 5.8e-202;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDPDSQPLNSLDVPLRKPRIPMETPRKVGIPITIALSLASIIIVVLKVIDKYYF 60
DB 1 MDPDSQPLNSLDVPLRKPRIPMETPRKVGIPITIALSLASIIIVVLKVIDKYYF 60
QY 61 LCGQPLHFIPIRQKLCDEBIDCPJGDEBHCVKSPFEGPAVAVRLSKDNSTLQVLSATGN 120
DB 61 LCGQPLHFIPIRQKLCDEBIDCPJGDEBHCVKSPFEGPAVAVRLSKDNSTLQVLSATGN 120
QY 121 WFSACFDNTEALAEATACROWGYSKPTPRAVEIGPDQDVVEITENSQELRMNSSGP 180
DB 121 WFSACFDNTEALAEATACROWGYSKPTPRAVEIGPDQDVVEITENSQELRMNSSGP 180
QY 181 CLSGSLVSLHCLACGSKSLKTPRVVGEBASVDSMPQVSIQYDKOHVCGSILDPHWVLT 240
DB 181 CLSGSLVSLHCLACGSKSLKTPRVVGEBASVDSMPQVSIQYDKOHVCGSILDPHWVLT 240
QY 241 AAHCRKHTDVFNWKVRAGSDKLGSPPSIAVAKIIIEFNPMYRKNDIALMKLQPLTF 300
DB 241 AAHCRKHTDVFNWKVRAGSDKLGSPPSIAVAKIIIEFNPMYRKNDIALMKLQPLTF 300
QY 301 SGTVRPICPFPEDELTPTPLMIIGWFTKONGKMSDILLQASVQVIDSTRCANADAY 360
DB 301 SGTVRPICPFPEDELTPTPLMIIGWFTKONGKMSDILLQASVQVIDSTRCANADAY 360
QY 361 QGEVTERKMKACGIPREGVDTCQDSDGSPLMYQSDQMHVGVISWGGCGSPSTPGVYTKV 420
DB 361 QGEVTERKMKACGIPREGVDTCQDSDGSPLMYQSDQMHVGVISWGGCGSPSTPGVYTKV 420
QY 421 SAYLNMWYVWKAEL 435
DB 421 SAYLNMWYVWKAEL 435

RESULT 4
ABU04912
ID ABU04912 standard; protein; 435 AA.
XX
AC ABU04912;
XX
DT 29-JAN-2003 (first entry)
XX
DE Human expressed protein tag (EPT) #1578.
XX
KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
KW protease; protease inhibitor; transporter; cytoskeletal protein;
KW receptor; transcription factor; cancer; MHC;
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200278524-A2.
XX
PD 10-OCT-2002.
XX
PF 28-MAR-2002; 2002WO-US009671.
XX
PR 28-MAR-2001; 2001US-0279495P.
PR 21-MAY-2001; 2001US-0292544P.
PR 08-AUG-2001; 2001US-0310801P.
```

CC and diagnosing diseases such as arthritis, chronic obstructive pulmonary
CC disorder (COPD), cancer, osteoporosis, aberrant wound healing,
CC angiogenesis, inflammatory disorders, diabetes, stroke and cardiovascular
CC diseases. Seripancrin genes are useful in chromosome localization
CC studies, as tools for tissue expression studies and also in gene therapy.
CC The polypeptides of the invention are used for identifying agonists and
CC antagonists useful for treating conditions associated with seripancrin
CC imbalance. These polypeptides are also useful as vaccines. The present
CC sequence is seripancrin protein. The seripancrin gene is located on human
CC chromosome 11q22-q23
XX
XX
SQ Sequence 435 AA;
Query Match 100.0%; Score 2342; DB 4; Length 435;
Best Local Similarity 100.0%; Pred. No. 5.8e-202;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDPSPDQPLNSLDVYKPKRPRIPIIALLSLASIIIVVLIKVLIDKYYF 60
Db 1 MDPSPDQPLNSLDVYKPKRPRIPIIALLSLASIIIVVLIKVLIDKYYF 60
QY 61 LCGQPLHPIPRKQICDGLDCLPGEDEBHCYKSPFEGPAVAVRLSKDRSTIQVDSATGN 120
Db 61 LCGQPLHPIPRKQICDGLDCLPGEDEBHCYKSPFEGPAVAVRLSKDRSTIQVDSATGN 120
QY 121 WFSACFDNFTALAEATACRQMGYSKPTFRAVEIGPDODLVVEITENSQELRKNSSGP 180
Db 121 WFSACFDNFTALAEATACRQMGYSKPTFRAVEIGPDODLVVEITENSQELRKNSSGP 180
QY 181 CLSGSLVSLHCLACGSKSLKTPRVVGGEBASVDSWPMQVSIQYDKQHVCGSILDPHWLT 240
Db 181 CLSGSLVSLHCLACGSKSLKTPRVVGGEBASVDSWPMQVSIQYDKQHVCGSILDPHWLT 240
QY 241 AAHCFRKATDVFNMKVRAGSDKLSGSPSLAVAKIIIEFNPMPYKNDIALMKIQPLTF 300
Db 241 AAHCFRKATDVFNMKVRAGSDKLSGSPSLAVAKIIIEFNPMPYKNDIALMKIQPLTF 300
QY 301 SGTVPRLPFPDELTATPLMTIIGMGFTKONGKMDIILQASVOYIIDSTRCANADAY 360
Db 301 SGTVPRLPFPDELTATPLMTIIGMGFTKONGKMDIILQASVOYIIDSTRCANADAY 360
QY 361 QGEVTEKMKACGIPREGVDTCCGDSGGLPMYOSDQMHVVGIVSMWGGCGSPGVTYKV 420
Db 361 QGEVTEKMKACGIPREGVDTCCGDSGGLPMYOSDQMHVVGIVSMWGGCGSPGVTYKV 420
QY 421 SAYLNMIVYVWKAEL 435
Db 421 SAYLNMIVYVWKAEL 435
RESULT 2
AAG78577 ID AAG78577 standard; protein; 435 AA.
XX
XX AAG78577;
AC
XX 07-MAY-2002 (first entry)
DT
XX
XX Human serine protease D-G amino acid sequence.
DE
XX
XX Serine protease; D-G; human; zymogen; enzyme; cytostatic;
KW antiinflammatory; dermatological; anticoagulation; cancer; skin disorder;
KW neuropathic pain; inflammatory disorder; coagulation diathesis;
KW thrombosis; laundry detergent; skin care; gene therapy.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH 31..52
FT Region /note="hydrophobic transmembrane domain"
FT Active-site 202..203
FT /note="residues spanning the zymogen cleavage site"
FT Active-site 243

FT /note= "catalytic triad residue"
FT Active-site 339
FT /note= "catalytic triad residue"
FT Misc-difference 361
FT /note= "Encoded by CTG"
FT Active-site 385
FT /note= "catalytic triad residue"
XX
XX WO200202011-A1.
XX
XX 10-JAN-2002.
XX
XX 08-JUN-2001; 2001WO-US018568.
XX
XX 30-JUN-2000; 2000US-00607745.
XX
XX (ORTH) ORTHO-MCNEIL PHARM INC.
XX
XX Darrow AL, Qi J, Andrade-Gordon P;
XX
XX WPI; 2002-106601/14.
XX
XX N-PSDB; AAI64284.
XX
XX Nucleic acid encoding a serine protease called D-G protein which is
XX useful for identifying modulators that are useful for treating a
XX condition which is mediated by protease D-G, e.g. cancer, skin disorders,
XX or neuropathic pain.
XX
XX Claim 13; Fig 1B; 81pp; English.
XX
XX The invention relates to an isolated and purified nucleic acid that
XX encodes a serine protease called D-G protein. The activity of the protein
XX of the invention may be described as cyostatic, antiinflammatory,
XX dermatological and anticoagulation. The serine protease of the invention
XX is a member of the trypsin/chymotrypsin-like (S1) serine protease family,
XX which play an important role in processes such as digestion and
XX regulatory amplification cascades through the proteolytic activation of
XX inactive zymogen precursors. Protease D-G modulating compounds are useful
XX for treating a condition which is mediated by protease D-G, e.g. cancer,
XX skin disorders, neuropathic pain, inflammatory disorders, or coagulation
XX diathesis/thrombosis. The polynucleotide encoding the protease is useful
XX for identifying, detecting or isolating mutant forms of DNA molecules
XX encoding the protease. The protease is useful for identifying modulators
XX of the functional protease. The D-G protein can be used for formulation
XX of compositions for laundry detergents and skin care products. Protease D
XX -G gene therapy may be used to introduce protease D-G into the cells of
XX target organisms. As the D-G protein is derived from a human, it is less
XX likely to produce an allergic reaction in sensitive individuals when used
XX in formulations for laundry detergents and skin care products. The
XX current sequence represents the human serine protease D-G amino acid
XX sequence
XX
XX Sequence 435 AA;
SQ
Query Match 100.0%; Score 2342; DB 5; Length 435;
Best Local Similarity 100.0%; Pred. No. 5.8e-202;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDPSPDQPLNSLDVYKPKRPRIPIIALLSLASIIIVVLIKVLIDKYYF 60
Db 1 MDPSPDQPLNSLDVYKPKRPRIPIIALLSLASIIIVVLIKVLIDKYYF 60
QY 61 LCGQPLHPIPRKQICDGLDCLPGEDEBHCYKSPFEGPAVAVRLSKDRSTIQVDSATGN 120
Db 61 LCGQPLHPIPRKQICDGLDCLPGEDEBHCYKSPFEGPAVAVRLSKDRSTIQVDSATGN 120
QY 121 WFSACFDNFTALAEATACRQMGYSKPTFRAVEIGPDODLVVEITENSQELRKNSSGP 180
Db 121 WFSACFDNFTALAEATACRQMGYSKPTFRAVEIGPDODLVVEITENSQELRKNSSGP 180
QY 181 CLSGSLVSLHCLACGSKSLKTPRVVGGEBASVDSWPMQVSIQYDKQHVCGSILDPHWLT 240
Db 181 CLSGSLVSLHCLACGSKSLKTPRVVGGEBASVDSWPMQVSIQYDKQHVCGSILDPHWLT 240

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2005, 13:41:18 ; Search time 188 Seconds

(without alignments)
1016.648 Million cell updates/sec

Title: US-10-803-530-2

Sequence: 1 MDPDSQPLNSLDVKEPLAKP.....VYTKVSAYLWYVWKAEL 435

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*
9: Geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2342	100.0	435	4 AAY72558	Aay72558 Human ser
2	2342	100.0	435	5 AAG78577	Aag78577 Human ser
3	2342	100.0	435	6 ABU04931	Abu04931 Human exp
4	2342	100.0	435	6 ABU04912	Abu04912 Human exp
5	2342	100.0	461	8 ADU91838	Adu91838 Ovr115. 2
6	2342	100.0	461	9 AEA10749	Aea10749 Mutated h
7	2338	99.8	435	2 AAY06437	Aay06437 Human pro
8	2338	99.8	435	6 ABU04930	Abu04930 Human exp
9	2337	99.8	437	5 AAG79359	Aag79359 CVA8 prei
10	2337	99.8	437	6 ABU04920	Abu04920 Human exp
11	2337	99.8	437	6 ABU04935	Abu04935 Human exp
12	2337	99.8	437	6 ABU04936	Abu04936 Human exp
13	2337	99.8	437	6 ABU04915	Abu04915 Human exp
14	2337	99.8	437	7 ADBR0525	Adbr0525 Human exp
15	2337	99.8	437	7 ADN39878	Adn39878 Ovarian c
16	2337	99.8	437	7 ADN39513	Adn39513 Cancer/an
17	2337	99.8	437	7 ADN39461	Adn39461 Cancer/an
18	2337	99.8	437	7 ADN39473	Adn39473 Cancer/an
19	2337	99.8	437	8 ADN04864	Adn04864 Antipsori
20	2337	99.8	437	8 ADU25740	Adu25740 Human TMP
21	2337	99.8	437	9 ADW04409	Adw04409 Human Ova
22	2337	99.8	437	9 AEA54849	Aea54849 Human tum
23	2333	99.6	435	8 ABO84561	Ab084561 Human can
24	2329	99.4	492	4 AAY72559	Aay72559 Human ser

25	2329	99.4	492	6 ABU04913	Abu04913 Human exp
26	2329	99.4	492	6 ABU04932	Abu04932 Human exp
27	2328	99.4	437	8 ABO84562	Ab084562 Human can
28	2319	99.0	437	4 AAE06931	Aae06931 Human mem
29	2319	99.0	437	6 ABU04917	Abu04917 Human exp
30	2319	99.0	437	6 ABU04922	Abu04922 Human exp
31	2319	99.0	437	7 AD110373	Ad110373 Human cel
32	2319	99.0	437	8 AD1346897	Ad1346897 Human tra
33	2297.5	98.1	432	3 AAY99417	Aay99417 Human tra
34	2297.5	98.1	432	4 AAB66166	Aab66166 Protein o
35	2297.5	98.1	432	4 AAB87581	Aab87581 Human PRO
36	2297.5	98.1	432	5 AAB87581	Abg95906 Human sec
37	2297.5	98.1	432	5 AAB87581	Aau76535 Tumour-as
38	2297.5	98.1	432	5 AAB87581	Abu58564 Human PRO
39	2297.5	98.1	432	6 ABU58564	Abu88112 Novel hum
40	2297.5	98.1	432	6 ABU84427	Abu84427 Human sec
41	2297.5	98.1	432	6 ABR66301	Abt66301 Human sec
42	2297.5	98.1	432	6 ABR66301	Abt66301 Human sec
43	2297.5	98.1	432	6 ABR66301	Abt66301 Human sec
44	2297.5	98.1	432	6 ABR66301	Abt66301 Human sec
45	2297.5	98.1	432	6 ABR66301	Abt66301 Human sec

ALIGNMENTS

RESULT 1
AAY72558
ID AAY72558 standard; protein, 435 AA.
AC AAY72558;
XX
DT 02-MAY-2001 (first entry)
XX
DE Human seripancrin protein.
XX
KW Human; seripancrin; serine protease; chromosome 11q22-q23; therapy;
KW arthralgia; chronic obstructive pulmonary disorder; COPD; cancer;
KW osteoporosis; aberrant wound healing; angiogenesis; diabetes;
KW inflammatory disorder; stroke; cardiovascular disease; gene therapy;
KW vaccine; cytostatic; cerebroprotective; vulnery; osteopathic.
XX
OS Homo sapiens.
XX
PN WO200104141-A2.
XX
PD 18-JAN-2001.
XX
PF 04-JUL-2000; 2000WO-EP006211.
XX
PR 12-JUL-1999; 99EP-00113428.
XX
PA (MERK) MERCK PATENT GMBH.
PI
PI Suendermann B, Hofmann U, Matzku S, Wilbert O;
XX
XX WPI; 2001-147177/15.
XX
DR N-PSDB; AAD02556.
XX
PT New extracellular serine protease Seripancrin, useful for treating
PT cancer, osteoporosis, arthritis, chronic obstructive pulmonary disease,
PT diabetes, inflammatory disorders, stroke, angiogenesis and aberrant wound
PT healing.
XX
PS Claim 2; Page 39-40; 45pp: English.
XX
CC The present invention relates to seripancrin polynucleotides, and
CC polypeptides encoded by them. Seripancrin are members of serine protease
CC family. This protein contains a transmembrane domain, a low density
CC lipoprotein (LDL) domain, protease domain and a scavenger receptor
CC cysteine-rich (SRCR) domain. The LDL and SRCR domains help to define the
CC specificity of seripancrin's intra and intermolecular interactions. The
CC polynucleotides and polypeptides of the invention are useful for treating

```
RESULT 14
US-09-851-588-6
; Sequence 6, Application US/09851588
; Patent No. 6682890
; ORGANISM: Homo sapiens
; GENERAL INFORMATION:
; APPLICANT: Mack, David
; APPLICANT: Gish, Kurt C.
; APPLICANT: Wilson, Keith E.
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSING COLORECTAL CANCER, COMPOSITIONS, AND
; TITLE OF INVENTION: OF SCREENING FOR COLORECTAL CANCER MODULATORS
; FILE REFERENCE: A-68829-1/DB/JUD/AMS
; CURRENT APPLICATION NUMBER: US/09/851,588
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 09/642,252
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: US 09/656,002
; PRIOR FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 406
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-851-588-6

Query Match          93.4%; Score 2188; DB 2; Length 406;
Best Local Similarity 100.0%; Pred. No. 6.1e-215;
Matches 406; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 VGPIPIIALLSLASIIIVVLIKVLIDKYFACGPHFIPRKQDCDELDCPLGEDEBH 89
DB 1 VGPIIALLSLASIIIVVLIKVLIDKYFACGPHFIPRKQDCDELDCPLGEDEBH 60

QY 90 CVKSPFEPGPAVAVRLSKDSTLOVLDSATGNMFACPDNFTALAEATACROWYSSKPTF 149
DB 61 CVKSPFEPGPAVAVRLSKDSTLOVLDSATGNMFACPDNFTALAEATACROWYSSKPTF 120

QY 150 RAVEIGPDODLVVEITENSQELMRNNSGPGCLSGSLVSLHCLACGSKLKTPTVVGGEEA 209
DB 121 RAVEIGPDODLVVEITENSQELMRNNSGPGCLSGSLVSLHCLACGSKLKTPTVVGGEEA 180

QY 210 SVDSNPMQVSTOYDQKHVCGGSIIDPHWVLTAAHCFRKHTDVFNKVRAGSDKLSFPSL 269
DB 181 SVDSNPMQVSTOYDQKHVCGGSIIDPHWVLTAAHCFRKHTDVFNKVRAGSDKLSFPSL 240

QY 270 AVAKIIIEFNPMYPRKNDIALMKLOPLTFSGTVRPICLPFDEELTPATPLMIIGMF 329
DB 241 AVAKIIIEFNPMYPRKNDIALMKLOPLTFSGTVRPICLPFDEELTPATPLMIIGMF 300

QY 330 TKONGGKMSDILLQASVQVYIDSTRCNADDAVQGEVTEKMCAGIPEGGVDTCCGDSGGPL 389
DB 301 TKONGGKMSDILLQASVQVYIDSTRCNADDAVQGEVTEKMCAGIPEGGVDTCCGDSGGPL 360

QY 390 MYQSDQMHVVGIVSWGYGCGSPSTPGVYTKVSAYLNMWYVWKAEI 435
DB 361 MYQSDQMHVVGIVSWGYGCGSPSTPGVYTKVSAYLNMWYVWKAEI 406

RESULT 15
US-09-607-745-9
; Sequence 9, Application US/09607745
; Patent No. 6750034
; GENERAL INFORMATION:
; APPLICANT: Darrow, Andrew L
; APPLICANT: Qi, Jain-shen
; APPLICANT: Andrade-Gordon, Patricia
; TITLE OF INVENTION: DNA encoding human serine protease D-G
; FILE REFERENCE: ORT-1273
; CURRENT APPLICATION NUMBER: US/09/607,745
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
```

```
; SEQ ID NO 9
; LENGTH: 292
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Fusion gene
US-09-607-745-9

Query Match          54.1%; Score 1266; DB 2; Length 292;
Best Local Similarity 97.4%; Pred. No. 7.1e-121;
Matches 228; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 202 RVVGGEEASVDSNPMQVSTOYDQKHVCGGSIIDPHWVLTAAHCFRKHTDVFNKVRAGSD 261
DB 51 KIVGGVALDVDSNPMQVSTOYDQKHVCGGSIIDPHWVLTAAHCFRKHTDVFNKVRAGSD 110

QY 262 KLSGFPFLAVAKIIIEFNPMYPRKNDIALMKLOPLTFSGTVRPICLPFDEELTPATP 321
DB 111 KLSGFPFLAVAKIIIEFNPMYPRKNDIALMKLOPLTFSGTVRPICLPFDEELTPATP 170

QY 322 LMIIGMGFTKONGGKMSDILLQASVQVYIDSTRCNADDAVQGEVTEKMCAGIPEGGVDTCC 381
DB 171 LMIIGMGFTKONGGKMSDILLQASVQVYIDSTRCNADDAVQGEVTEKMCAGIPEGGVDTCC 230

QY 382 QGDSGGPLMTQSDQMHVVGIVSWGYGCGSPSTPGVYTKVSAYLNMWYVWKAEI 435
DB 231 QGDSGGPLMTQSDQMHVVGIVSWGYGCGSPSTPGVYTKVSAYLNMWYVWKAEI 284

Search completed: December 5, 2005, 13:54:42
Job time : 47 secs
```

Query Match	98.1%	Score 2297.5	DB 2	Length 433
Best Local Similarity	98.8%	Pred. No. 4,3e-226		
Matches	429	Conservative	0	Mismatches 0
			Indels	5
			Gaps	1
QY	2	DPDSQPLNSLDVKELRKPRIPMETFRKGIPIIIALLSLASIIIVVLIKYILDKYFL	61	
Db	4	DPDSQPLNSLDVKELRKPRIPMETFRKGIPIIIALLSLASIIIVVLIKYILDKYFL	63	
QY	62	CGQPLHPIRQQLCGELDCPLGEDBEHCVKSFPPEGPAVAFLSKDRSTLQVLDSATGNW	121	
Db	64	CGQPLHPIRQQLCGELDCPLGEDBEHCVKSFPPEGPAVAFLSKDRSTLQVLDSATGNW	123	
QY	122	FSACEPDNTEALAEACRQMGYSKPTTFAVEIGPDQDLVEITENSGELMRNMSGPC	181	
Db	124	FSACEPDNTEALAEACRQMGYS-----RAVEIGPDQDLVEITENSGELMRNMSGPC	178	
QY	182	LSSGLVSLHCLACGSLKTPRVVGGEEASVDSMPQVSIQYDKQVCGGSIIDPHVTLTA	241	
Db	179	LSSGLVSLHCLACGSLKTPRVVGGEEASVDSMPQVSIQYDKQVCGGSIIDPHVTLTA	238	
QY	242	AHCRKRDVENVKVRAGSDKLGSPPSLAVKIIIIIEENPMVKNDIALMLQGPLPLFS	301	
Db	239	AHCRKRDVENVKVRAGSDKLGSPPSLAVKIIIIIEENPMVKNDIALMLQGPLPLFS	298	

Query	March	93.8%	Score	2196	DB 2	Length	423
Best Local Similarity	97.1%	Pred. No.	9.9e-216				
Matches	408	Conservative	1	Mismatches	11	Indels	0
							Gaps
							0
QY	16	PLKRPRIEMETFRKVGPIPIIALSLASIIIVVLLIKVILDKYFLCGQPLHIFIRKOLC	75				
Db	4	PCANPVPMPWRSESVGPIPIIALSLASIIIVVLLIKVILDKYFLCGQPLHIFIRKOLC	63				
QY	76	DGEILDCLGDBDEHCVKVSFPBGPAAVAVRLSKDRSTLOYLSATGNWFSACDNTFEALAE	135				
Db	64	DGEILDCLGDBDEHCVKVSFPBGPAAVAVRLSKDRSTLOYLSATGNWFSACDNTFEALAE	123				
QY	136	TACRQMGVSSKPTFRAYEIGPDODLDVETIENSQELRMRRSSGBCJLSSIVSLHCLACG	195				
Db	124	TACRQMGVSSKPTFRAYEIGPDODLDVETIENSQELRMRRSSGBCJLSSIVSLHCLACG	183				
QY	196	KSILKTPRVVGGEEASVDSWPQVSIQYDKQVCGSILDPHMYLTAACFRKHTDVFNMK	255				
Db	184	KSILKTPRVVGGEEASVDSWPQVSIQYDKQVCGSILDPHMYLTAACFRKHTDVFNMK	243				
QY	256	VRASDCLGSPPLAAVAKIIIEFNPMYPRDNDIALMKLOPLPFGSGVRIICLPFPEE	315				
Db	244	VRASDCLGSPPLAAVAKIIIEFNPMYPRDNDIALMKLOPLPFGSGVRIICLPFPEE	303				
QY	316	LTPATPLMIIGWGFTKNGGKMSDILLQASQVUIDSTRCNADDAAYOGSEVTEKMKCAGIPE	375				
Db	304	LTPATPLMIIGWGFTKNGGKMSDILLQASQVUIDSTRCNADDAAYOGSEVTEKMKCAGIPE	363				
QY	376	GGVDTCGGDSGGPLMYOSDQMHVVGIVSKWGGCGSPSTPGVYTKVASIYMLIYVWKAEL	435				
Db	364	GGVDTCGGDSGGPLMYOSDQMHVVGIVSKWGGCGSPSTPGVYTKVASIYMLIYVWKAEL	423				

QY 422 AYLNWIYNVWKAEL 435
 |||||
Db 419 AYLNWIIYNVWKAEL 432

```

RESULT 10
US-10-011-833A-275
; Sequence 275, Application US/10011833A
; Patent No. 6951920
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan I.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2830PIC22
; TITLE OF INVENTION: Acids Encoding the Same
; CURRENT APPLICATION NUMBER: US/10/011,833A
; PRIOR FILING DATE: 2002-06-25
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 275
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-011-833A-275

```

Query Match	98.1%	Score 2297.5	DB 2	Length 432
Best Local Similarity	98.8%	Pred. No. 4,3e-26;		
Matches 429;	Conservative 0;	Matches 0;	Indels 5;	Gaps 14;
QY	2	DDSDQPLNSLDVKELARKPRIPMEFFRKAGPIITALLSLASIIIVVLIKILDKXYEL	61	
DB	4	DDSDQPLNSLDVKELARKPRIPMEFFRKAGPIITALLSLASIIIVVLIKILDKXYFL	63	
QY	62	CGQPLHFIPIRKOLCGEILDCLPGEDBEHCYKSPFEGPAVAAYLSKDRSTLYLDSATGNW	121	
DB	64	CGQPLHFIPIRKOLCGEILDCLPGEDBEHCYKSPFEGPAVAAYLSKDRSTLYLDSATGNW	123	
QY	122	FSACPDNFTLEALAEACROMGYSSKPTFAVEIGPDODLVEIITENSQELMRNSSGPC	181	
DB	124	FSACPDNFTLEALAEACROMGYS-----RAVEIGPDODLVEIITENSQELMRNSSGPC	178	
QY	182	LSGLSVLSLHCLACGSKSLKTPRVVGGEEASVDSWPQVSIQYDKOHVCGGSIIDPHWVLT	241	
DB	179	LSGLSVLSLHCLACGSKSLKTPRVVGGEEASVDSWPQVSIQYDKOHVCGGSIIDPHWVLT	238	
QY	242	AHCFKPHDVFWMKTRAGSDKLGSPSLAVAKIIIIIEFPMPKPNODIALMLQLQPLTFS	301	
DB	239	AHCFKPHDVFWMKTRAGSDKLGSPSPSLAVAKIIIIIEFPMPKPNODIALMLQLQPLTFS	298	
QY	302	GVVRPICLPFPEDELTPTAPLMIIGMGFTKONGGKMSDILQASVOVIDSTRCANDDAYQ	361	
DB	299	GVVRPICLPFPEDELTPTAPLMIIGMGFTKONGGKMSDILQASVOVIDSTRCANDDAYQ	358	
QY	362	GEVTERKMCAGIPBEGVDTCQGDGGPLMYOSDQMHVVGIVISWGYCGCGPSTPGVYTKVS	421	
DB	359	GEVTERKMCAGIPBEGVDTCQGDGGPLMYQSDQMHVVGIVISWGYCGCGPSTPGVYTKVS	418	
QY	422	AYLNNIYVWKAEI	435	
DB	419	AYLNNIYVWKAEI	432	

RESULT 11
US-10-006-041A-275
; Sequence 275, Application US/10006041A

```

/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Botstein, David
/ APPLICANT: Deenoyers, Luc
/ APPLICANT: Eaton, Dan I.
/ APPLICANT: Ferrara, Napoleone
/ APPLICANT: Fong, Sherman
/ APPLICANT: Gao, Wei-Qiang
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Grimaldi, Christopher J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Hillan, Kenneth J.
/ APPLICANT: Pan, James
/ APPLICANT: Paoni, Nicholas F.
/ TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
/ TITLE OF INVENTION: Acids Encoding the Same
/ FILE REFERENCE: P2830P18
/ CURRENT APPLICATION NUMBER: US/10/006,041A
/ CURRENT FILING DATE: 2001-12-06
/ Prior Application removed - See File Wrapper or Palm
/ NUMBER OF SEQ ID NOS: 477
/ SEQ ID NO 275
/ LENGTH: 432
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-006-041A-275

```

Query Match	98.1%	Score 2297.5	DB 2	Length 432
Best Local Similarity	98.8%	Pred. No. 4.3e-226		
Matches 429	Conservative 0	Mismatches 0	Indels 5	Gaps 1
Qy	2	DPDSQPLNSLDVPLRPRIPIMETFRKVGIPITIALSLASIIIVVLLIRVILDKYFL	61	
Db	4	DPDSQPLNSLDVPLRPRIPIMETFRKVGIPITIALSLASIIIVVLLIRVILDKYFL	63	
Qy	62	CGQPLHPIPRKQLCDGELDCPLGDEDEHCVKSPEGPAAVAVRLSKDRSTLQVLSATGNW	121	
Db	64	CGQPLHPIPRKQLCDGELDCPLGDEDEHCVKSPEGPAAVAVRLSKDRSTLQVLSATGNW	123	
Qy	122	FSACDNTTEALAEACRQNGYSKAPTRAVETIGPDQDLVDVETITENSQELRMRSSGPC	181	
Db	124	FSACDNTTEALAEACRQNGYSKAPTRAVETIGPDQDLVDVETITENSQELRMRSSGPC	178	
Qy	182	LSGSLVSLHCLACGSKLTPTRVVGESEASVSWMPQVSIQYDKQHCVCSSILDPHMYLTA	241	
Db	179	LSGSLVSLHCLACGSKLTPTRVVGESEASVSWMPQVSIQYDKQHCVCSSILDPHMYLTA	238	
Qy	242	AHCFKHTDVFNMKVRAGSDKLGSPSLAVAKIIIEENPMYPKNDITLMLKLOPLTFSS	301	
Db	239	AHCFKHTDVFNMKVRAGSDKLGSPSLAVAKIIIEENPMYPKNDITLMLKLOPLTFSS	298	
Qy	302	GTVPDICLPFDEEITPATPLMIIGWGTTKONGGMSDILLQASVQVITDSTRCNADDAVQ	361	
Db	299	GTVPDICLPFDEEITPATPLMIIGWGTTKONGGMSDILLQASVQVITDSTRCNADDAVQ	358	
Qy	362	GEVTEKMKCAGIPREGVPTCCGDSGGRIMYOSDQMHVVGIVSWGCGGSPSPGYYTKVS	421	
Db	359	GEVTEKMKCAGIPREGVPTCCGDSGGRIMYOSDQMHVVGIVSWGCGGSPSPGYYTKVS	418	
Qy	422	AYLNIYVWKAEL	435	
Db	419	AYLNIYVWKAEL	432	

Qy	302	GTVRICLPFPEBELTPATPLMIIGMFTKQNGKKSDILLQASVOVIDSTCNDDAAYQ	361
Db	299	GTVRICLPFPEBELTPATPLMIIGMFTKQNGKKSDILLQASVOVIDSTCNDDAAYQ	358
Qy	362	GEVTERKMKACIGPEGVDTPCGDSGGPLMYQSDQWHVGI VSMGYCGGSPSTPGVYTKVS	421
Db	359	GEVTERKMKACIGPEGVDTPCGDSGGPLMYQSDQWHVGI VSMGYCGGSPSTPGVYTKVS	418
Qy	422	AYLNIWYVWKAEI 435	
Db	419	AYLNIWYVWKAEI 432	
RESULT 8			
US-10-015-671A-275			
; Sequence 275, Application US/10015671A			
; Patent No. 6946263			
; GENERAL INFORMATION:			
; APPLICANT: Baker, Kevin P.			
; APPLICANT: Botstein, David			
; APPLICANT: Desnoyers, Luc			
; APPLICANT: Eaton, Dan I.			
; APPLICANT: Ferrara, Napoleone			
; APPLICANT: Fong, Sherman			
; APPLICANT: Gao, Wei-Qiang			
; APPLICANT: Goddard, Audrey			
; APPLICANT: Grimaldi, Paul J.			
; APPLICANT: Gurney, Austin J.			
; APPLICANT: Hillan, Kenneth J.			
; APPLICANT: Pan, James			
; APPLICANT: Paoni, Nicholas F.			
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic			
; TITLE OF INVENTION: Acids Encoding the Same			
; FILE REFERENCE: P2830PLC47			
; CURRENT APPLICATION NUMBER: US/10/015,671A			
; PRIOR FILING DATE: 2001-12-11			
; Prior application removed - See File Wrapper or Palm			
; NUMBER OF SEQ ID NOS: 477			
; SEQ ID NO 275			
; LENGTH: 432			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-10-015-671A-275			
Query Match 98.1%; Score 2297.5; DB 2; Length 432;			
Best Local Similarity 98.8%; Pred. No. 4,3e-226;			
Matches 429; Conservative 0; Mismatches 26; Indels 5; Gaps 1			
Qy	2	DPDSQPLNSLDVKRLRKPRIMEFTRKVGIPILIALSLASIIIVVLIKYILDKRYFL	61
Db	4	DPDSQPLNSLDVKRLRKPRIMEFTRKVGIPILIALSLASIIIVVLIKYILDKRYFL	63
Qy	62	CGQPLHFI PRKQLCDGELDCPLGEDEEHCVKSPFEGPAVAVRLSKDRSTLQVLD SATGNW	121
Db	64	CGQPLHFI PRKQLCDGELDCPLGEDEEHCVKSPFEGPAVAVRLSKDRSTLQVLD SATGNW	123
Qy	122	FSACDNTTEALAEFAACRQMGVSSKPTPAVAITGPDQDLDVETITENSGELAMRNSGGC	181
Db	124	FSACDNTTEALAEFAACRQMGVSSKPTPAVAITGPDQDLDVETITENSGELAMRNSGGC	178
Qy	182	LSGSLVSLHCLACGSLKTPRVVVGEEASVDSMPQVSIQYKQNVCGSILDPHVLTA	241
Db	179	LSGSLVSLHCLACGSLKTPRVVVGEEASVDSMPQVSIQYKQNVCGSILDPHVLTA	238
Qy	242	AHCFKHTDVFVWKVRAGSDKLGSPSLAVAKIIIEFNPMYKXNDIALMKLQEP LTF	301
Db	239	AHCFKHTDVFVWKVRAGSDKLGSPSLAVAKIIIEFNPMYKXNDIALMKLQEP LTF	298
Qy	302	GTVRICLPFPEBELTPATPLMIIGMFTKQNGKKSDILLQASVOVIDSTCNDDAAYQ	361
Db	299	GTVRICLPFPEBELTPATPLMIIGMFTKQNGKKSDILLQASVOVIDSTCNDDAAYQ	358

QY	362	GEVEEKMKACGAIPEGSGVDTCCGDSGGPRLMYOSDOMHVVGIYSWKGCGGSPSTPGVYTKVS	421
Db	359	GEVEEKMKACGAIPEGSGVDTCCGDSGGPRLMYOSDOMHVVGIYSWKGCGGSPSTPGVYTKVS	418
QY	422	AYLWMIYVWKAEL 435	
Db	419	AYLWMIYVWKAEL 432	

RESULT 9
US-10-015-393A-275
Sequence 275, Application US/10015393A
Patent No. 6951737
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Deonoyers, Luc
APPLICANT: Eaton, Dan I.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2830P1C46
CURRENT FILING DATE: 2002-06-10
Prior Application removed - See File wrapper or Palm
SEQ ID NO 275
NUMBER OF SEQ ID NOS: 477
LENGTH: 432
TYPE: PRT
ORGANISM: Homo sapiens
US-10-015-393A-275

Query Match	98.1%;	Score 2297.5;	DB 2;	Length 432;
Best Local Similarity	98.8%;	Pred. No. 4.3e-226;		
Matches 429;	Conservative 0;	Mismatches 0;	Indels 5;	Gaps 1

QY	2	DPDSDDPLNSIDVPLRKPRIPMETFRKVGIPITIIALISLASSIITVVLIKVIIDKYYFL	61
Db	4	DPDSDDPLNSIDVPLRKPRIPMETFRKVGIPITIIALISLASSIITVVLIKVIIDKYYFL	63
QY	62	CGQPLHPIPRKQOLDGELDCPLGDEDEHCVSPFEGPAVAARLSKDRSTLOVLSATGNW	121
Db	64	CGQPLHPIPRKQOLDGELDCPLGDEDEHCVSPFEGPAVAARLSKDRSTLOVLSATGNW	123
QY	122	FSACFDNFTEALAEATACRQMGYSKPTFRAYEIPDDDLDAVEITENSOELMRNNSGPGC	181
Db	124	FSACFDNFTEALAEATACRQMGYS-----RAYEIPDDDLDAVEITENSOELMRNNSGPGC	178
QY	182	LSSGLVSLHCLACGSKLTPRVVVGGEBAVSVDSPWQVSIQYDKOHVCGGSLIDPHMVLTA	241
Db	179	LSSGLVSLHCLACGSKLTPRVVVGGEBAVSVDSPWQVSIQYDKOHVCGGSLIDPHMVLTA	238
QY	242	AHCRKKTIDVNMKVARGSDGLGSPSLAVAKIIITFENPMYPKNDIDALMKLOPLTFSS	301
Db	239	AHCRKKTIDVNMKVARGSDGLGSPSLAVAKIIITFENPMYPKNDIDALMKLOPLTFSS	298
QY	302	GTVPAPICLPFDEBELTPATPLMIIGWGFTHKONGGKMSDIILOQASVOVIDSTRCNADAYQ	361
Db	299	GTVPAPICLPFDEBELTPATPLMIIGWGFTHKONGGKMSDIILOQASVOVIDSRCNADAYQ	358
QY	362	GEVEEKMKACGAIPEGSGVDTCCGDSGGPRLMYOSDOMHVVGIYSWKGCGGSPSTPGVYTKVS	421
Db	359	GEVEEKMKACGAIPEGSGVDTCCGDSGGPRLMYOSDOMHVVGIYSWKGCGGSPSTPGVYTKVS	418

QY 182 LSGSLVSIHCLACGKSLKTPRVVGGEEASVDSWPMQVSIQYDKQHVCGGSLIDPHWVLT 241
DB 179 LSGSLVSIHCLACGKSLKTPRVVGGEEASVDSWPMQVSIQYDKQHVCGGSLIDPHWVLT 238
QY 242 AHCFRKHTDVFNWVKVRAGSDKLSFSPSLAVAKIIIEFNPMYPRXNDIALMKLOFPLTFS 301
DB 239 AHCFRKHTDVFNWVKVRAGSDKLSFSPSLAVAKIIIEFNPMYPRXNDIALMKLOFPLTFS 298
QY 302 GTVRPCLPFDEBELTPATPLMIIGMFTKONGKMSDILLQASVOYIDSTRCANADAYQ 361
DB 299 GTVRPCLPFDEBELTPATPLMIIGMFTKONGKMSDILLQASVOYIDSTRCANADAYQ 358
QY 362 GEVTERKMKAGIPBGGVDTCCGDSGGPLMYQSDQMHVGVISWGYCGGSPSTPGVYTKVS 421
DB 359 GEVTERKMKAGIPBGGVDTCCGDSGGPLMYQSDQMHVGVISWGYCGGSPSTPGVYTKVS 418
QY 422 AYLMWYVWKAEL 435
DB 419 AYLMWYVWKAEL 432

RESULT 6

US-10-015-389A-275
; Sequence 275, Application US/10015389A
; Patent No. 6936436
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Deenoyers, Luc
; APPLICANT: Eaton, Dan I.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2830P1C48
; CURRENT APPLICATION NUMBER: US/10/015,389A
; PRIORITY FILING DATE: 2002-06-25
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 275
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-015-389A-275

Query Match 98.1%; Score 2297.5; DB 2; Length 432;
Best Local Similarity 98.8%; Pred. No. 4.3e-226;
Matches 429; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY 2 DPDSQDPLNSLDVPLRKPRIPMETFRKVGIPITIALSLASIIIVVLLIKVILDKYYFL 61
DB 4 DPDSQDPLNSLDVPLRKPRIPMETFRKVGIPITIALSLASIIIVVLLIKVILDKYYFL 63
QY 62 CGQPLHFIIPKQCLDGEIDCPLGDEBEHCYKSPFEGPAVAVRLSKDSTLQVLDATGNW 121
DB 64 CGQPLHFIIPKQCLDGEIDCPLGDEBEHCYKSPFEGPAVAVRLSKDSTLQVLDATGNW 123
QY 122 FSAFDFNTEALTAETACRQMGYSKPTFRAVEIGPDODLVETENSOELRMNNSGPC 181
DB 124 FSAFDFNTEALTAETACRQMGYSKPTFRAVEIGPDODLVETENSOELRMNNSGPC 178
QY 182 LSGSLVSIHCLACGKSLKTPRVVGGEEASVDSWPMQVSIQYDKQHVCGGSLIDPHWVLT 241
DB 179 LSGSLVSIHCLACGKSLKTPRVVGGEEASVDSWPMQVSIQYDKQHVCGGSLIDPHWVLT 238

QY 242 AHCFRKHTDVFNWVKVRAGSDKLSFSPSLAVAKIIIEFNPMYPRXNDIALMKLOFPLTFS 301
DB 239 AHCFRKHTDVFNWVKVRAGSDKLSFSPSLAVAKIIIEFNPMYPRXNDIALMKLOFPLTFS 298
QY 302 GTVRPCLPFDEBELTPATPLMIIGMFTKONGKMSDILLQASVOYIDSTRCANADAYQ 361
DB 299 GTVRPCLPFDEBELTPATPLMIIGMFTKONGKMSDILLQASVOYIDSTRCANADAYQ 358
QY 362 GEVTERKMKAGIPBGGVDTCCGDSGGPLMYQSDQMHVGVISWGYCGGSPSTPGVYTKVS 421
DB 359 GEVTERKMKAGIPBGGVDTCCGDSGGPLMYQSDQMHVGVISWGYCGGSPSTPGVYTKVS 418
QY 422 AYLMWYVWKAEL 435
DB 419 AYLMWYVWKAEL 432

RESULT 7

US-10-006-768A-275
; Sequence 275, Application US/10006768A
; Patent No. 6936697
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Deenoyers, Luc
; APPLICANT: Eaton, Dan I.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2830P1C10
; CURRENT APPLICATION NUMBER: US/10/006,768A
; PRIORITY FILING DATE: 2002-03-05
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 275
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-006-768A-275

Query Match 98.1%; Score 2297.5; DB 2; Length 432;
Best Local Similarity 98.8%; Pred. No. 4.3e-226;
Matches 429; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY 2 DPDSQDPLNSLDVPLRKPRIPMETFRKVGIPITIALSLASIIIVVLLIKVILDKYYFL 61
DB 4 DPDSQDPLNSLDVPLRKPRIPMETFRKVGIPITIALSLASIIIVVLLIKVILDKYYFL 63
QY 62 CGQPLHFIIPKQCLDGEIDCPLGDEBEHCYKSPFEGPAVAVRLSKDSTLQVLDATGNW 121
DB 64 CGQPLHFIIPKQCLDGEIDCPLGDEBEHCYKSPFEGPAVAVRLSKDSTLQVLDATGNW 123
QY 122 FSAFDFNTEALTAETACRQMGYSKPTFRAVEIGPDODLVETENSOELRMNNSGPC 181
DB 124 FSAFDFNTEALTAETACRQMGYSKPTFRAVEIGPDODLVETENSOELRMNNSGPC 178
QY 182 LSGSLVSIHCLACGKSLKTPRVVGGEEASVDSWPMQVSIQYDKQHVCGGSLIDPHWVLT 241
DB 179 LSGSLVSIHCLACGKSLKTPRVVGGEEASVDSWPMQVSIQYDKQHVCGGSLIDPHWVLT 238
QY 242 AHCFRKHTDVFNWVKVRAGSDKLSFSPSLAVAKIIIEFNPMYPRXNDIALMKLOFPLTFS 301
DB 239 AHCFRKHTDVFNWVKVRAGSDKLSFSPSLAVAKIIIEFNPMYPRXNDIALMKLOFPLTFS 298

```
Query Match          99.8%; Score 2338; DB 2; Length 435;
Best Local Similarity 99.8%; Pred. No. 3.1e-230;
Matches 434; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPDSQPLNSLDVKKLRKPRIPMETFRKVGIPITIALSLASIIIVVLIKYLIDKYYFL 60
DB 1 MDPDSQPLNSLDVKKLRKPRIPMETFRKVGIPITIALSLASIIIVVLIKYLIDKYYFL 60
QY 61 LCGQPLHFIIPRKQOLCGELDCPLGDEDEHCVKSPFEGPAVAVRLSKDRSTLQVLDSATGN 120
DB 61 LCGQPLHFIIPRKQOLCGELDCPLGDEDEHCVKSPFEGPAVAVRLSKDRSTLQVLDSATGN 120
QY 121 WPSACDNFTLEALAEACRQMGYSKPTFRAYEIGPDQDLVVEITENSQELMRNSSGP 180
DB 121 WPSACDNFTLEALAEACRQMGYSKPTFRAYEIGPDQDLVVEITENSQELMRNSSGP 180
QY 121 WPSACDNFTLEALAEACRQMGYSKPTFRAYEIGPDQDLVVEITENSQELMRNSSGP 180
DB 121 WPSACDNFTLEALAEACRQMGYSKPTFRAYEIGPDQDLVVEITENSQELMRNSSGP 180
QY 181 CLSGSLVSLHCLACGSLKTPRVVGESEASVDSMPQVSIQYDKQVCCGSIIDPHWVLT 240
DB 181 CLSGSLVSLHCLACGSLKTPRVVGESEASVDSMPQVSIQYDKQVCCGSIIDPHWVLT 240
QY 241 AAHCFRKHIDVFEMKVRAGSDKLGSPSLAVAKIIIEFNPMYPKNDIALMKLQPLTF 300
DB 241 AAHCFRKHIDVFEMKVRAGSDKLGSPSLAVAKIIIEFNPMYPKNDIALMKLQPLTF 300
QY 301 SGTVRPCLPFDDEELTPATPLMIIGWFTKQNGKMSDILLQASVQVIDSTRCNADAY 360
DB 301 SGTVRPCLPFDDEELTPATPLMIIGWFTKQNGKMSDILLQASVQVIDSTRCNADAY 360
QY 361 QGEVTERKMCAGIPFEGGVDTCCGDSGGLMYQSDQMHVVGIVSMGCGGSPSTPGVYTV 420
DB 361 QGEVTERKMCAGIPFEGGVDTCCGDSGGLMYQSDQMHVVGIVSMGCGGSPSTPGVYTV 420
QY 421 SAYLNIYVWKAEI 435
DB 421 SAYLNIYVWKAEI 435

RESULT 4
US-09-851-588-8
; Sequence 8, Application US/09851588
; Patent No. 6682890
; GENERAL INFORMATION:
; APPLICANT: Mack, David
; APPLICANT: Gish, Kurt C.
; APPLICANT: Wilson, Keith E.
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSING COLORECTAL CANCER, COMPOSITIONS, AND
; FILE REFERENCE: A-68829-1/DJL/JJD/AMS
; CURRENT APPLICATION NUMBER: US/09/851,588
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 09/642,252
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: US 09/656,002
; PRIOR FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-851-588-8

Query Match          99.8%; Score 2337; DB 2; Length 437;
Best Local Similarity 100.0%; Pred. No. 4e-230;
Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 122 FSACDNFTLEALAEACRQMGYSKPTFRAYEIGPDQDLVVEITENSQELMRNSSGPC 181
DB 124 FSACDNFTLEALAEACRQMGYSKPTFRAYEIGPDQDLVVEITENSQELMRNSSGPC 183
QY 182 LSGSLVSLHCLACGSLKTPRVVGESEASVDSMPQVSIQYDKQVCCGSIIDPHWVLT 241
DB 184 LSGSLVSLHCLACGSLKTPRVVGESEASVDSMPQVSIQYDKQVCCGSIIDPHWVLT 243
QY 242 AAHCFRKHIDVFEMKVRAGSDKLGSPSLAVAKIIIEFNPMYPKNDIALMKLQPLTF 301
DB 244 AAHCFRKHIDVFEMKVRAGSDKLGSPSLAVAKIIIEFNPMYPKNDIALMKLQPLTF 303
QY 302 GTVRPCLPFDDEELTPATPLMIIGWFTKQNGKMSDILLQASVQVIDSTRCNADAY 361
DB 304 GTVRPCLPFDDEELTPATPLMIIGWFTKQNGKMSDILLQASVQVIDSTRCNADAY 363
QY 362 GEVTERKMCAGIPFEGGVDTCCGDSGGLMYQSDQMHVVGIVSMGCGGSPSTPGVYTV 421
DB 364 GEVTERKMCAGIPFEGGVDTCCGDSGGLMYQSDQMHVVGIVSMGCGGSPSTPGVYTV 423
QY 422 SAYLNIYVWKAEI 435
DB 424 SAYLNIYVWKAEI 437

RESULT 5
US-10-012-231A-275
; Sequence 275, Application US/10012231A
; Patent No. 6924355
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Baton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoli, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P28301C23
; CURRENT APPLICATION NUMBER: US/10/012,231A
; PRIOR FILING DATE: 2002-06-10
; PRIOR Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 275
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-012-231A-275

Query Match          98.1%; Score 2297.5; DB 2; Length 432;
Best Local Similarity 98.8%; Pred. No. 4.3e-226;
Matches 429; Conservative 0; Mismatches 0; Indels 5; Gaps 1;
```

Db 361 QGEVTERKMCAGIPGEGVDTCQSDSGGPLYMQSDQMHWVGIWSMGYCGGSPSTPGVYTKV 420
QY 421 SAYLNMWYWKAEI 435
Db 421 SAYLNMWYWKAEI 435

RESULT 2

US-09-008-271A-6

; Sequence 6, Application US/09008271A
; Patent No. 6203979
; GENERAL INFORMATION:

APPLICANT: Bandman, Olga
; Hillman, Jennifer L.
; Yue, Henry
; Guegler, Karl J.
; Corley, Neil C.
; Tang, Tom Y.
; Shah, Purvi

TITLE OF INVENTION: HUMAN PROTEASE MOLECULES
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/008,271A
FILING DATE: 16-Jan-1998

PRIOR APPLICATION DATA:
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: Mohan-Peterson, Sheela
REGISTRATION NUMBER: 41,201
REFERENCE/DOCKET NUMBER: PF-0458 US

TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 435 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: COLNOT13

CLONE: 1337018

SEQUENCE DESCRIPTION: SEQ ID NO: 6

Query Match 99.8%; Score 2338; DB 2; Length 435;
Best Local Similarity 99.8%; Pred. No. 3,1e-230;

Matches 434; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPDSQPLNSIDVPLKPRIPMETFRKVGPIIIALLSLASIIIVVLLKVIIDKYYF 60
Db 1 MDPDSQPLNSIDVPLKPRIPMETFRKVGPIIIALLSLASIIIVVLLKVIIDKYYF 60
QY 61 LCGQPLAIFPRKQLCDGLDCLGDEDEHCVKSPFEGPAVAVRLSKDRSTIQVLDSATGN 120
Db 61 LCGQPLAIFPRKQLCDGLDCLGDEDEHCVKSPFEGPAVAVRLSKDRSTIQVLDSATGN 120
QY 121 WFSACFDNFTEALATACRQMGYSKPTFRFAVEIPDDLDVVEITENSQEIEMNSSGP 180
Db 121 WFSACFDNFTEALATACRQMGYSKPTFRFAVEIPDDLDVVEITENSQEIEMNSSGP 180

QY 181 CLSGSLVSLHCLACGSKSLKTPRVVGCSEASVDSWPMQVSIQYDKQHYCGGSILDPHWLT 240
Db 181 CLSGSLVSLHCLACGSKSLKTPRVVGCSEASVDSWPMQVSIQYDKQHYCGGSILDPHWLT 240
QY 241 AAFGRKHTDVFNKVRAGSDKLGSPSLAVAKIIIEFNPMYRKNDIALMKIQPLTF 300
Db 241 AAFGRKHTDVFNKVRAGSDKLGSPSLAVAKIIIEFNPMYRKNDIALMKIQPLTF 300
QY 301 SGTVPICLPPEDELTPATPMTITGFTKONGCKMSDILLOASVOYIDSTRCNADAY 360
Db 301 SGTVPICLPPEDELTPATPMTITGFTKONGCKMSDILLOASVOYIDSTRCNADAY 360
QY 361 QGEVTERKMCAGIPGEGVDTCQSDSGGPLYMQSDQMHWVGIWSMGYCGGSPSTPGVYTKV 420
Db 361 QGEVTERKMCAGIPGEGVDTCQSDSGGPLYMQSDQMHWVGIWSMGYCGGSPSTPGVYTKV 420
QY 421 SAYLNMWYWKAEI 435
Db 421 SAYLNMWYWKAEI 435

RESULT 3

US-09-968-415-6

; Sequence 6, Application US/09968415
; Patent No. 6855811
; GENERAL INFORMATION:

APPLICANT: Bandman, Olga
; Hillman, Jennifer L.
; Yue, Henry
; Guegler, Karl J.
; Corley, Neil C.
; Tang, Tom Y.
; Shah, Purvi

TITLE OF INVENTION: HUMAN PROTEASE MOLECULES
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/968,415
FILING DATE: 26-Sep-2001

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/659,151
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: Mohan-Peterson, Sheela
REGISTRATION NUMBER: 41,201
REFERENCE/DOCKET NUMBER: PF-0458 US

TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 435 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: COLNOT13

SEQUENCE DESCRIPTION: SEQ ID NO: 6

US-09-968-415-6

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2005, 13:45:44 ; Search time 45 Seconds

(Without alignments)
799.198 Million cell updates/sec

Title: US-10-803-530-2

Perfect score: 2342

Sequence: 1 MDPDSQPLNSLDVFKLRKP.....VYTKSAYLNMIYWKAEI 435

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/iaa/5.COMB.pep.*
2: /cgn2_6/prodata/1/iaa/6.COMB.pep.*
3: /cgn2_6/prodata/1/iaa/H.COMB.pep.*
4: /cgn2_6/prodata/1/iaa/PCTUS.COMB.pep.*
5: /cgn2_6/prodata/1/iaa/R.COMB.pep.*
6: /cgn2_6/prodata/1/iaa/backfilese1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2342	100.0	435	2	US-09-607-745-2
2	2338	99.8	435	2	US-09-008-271A-6
3	2338	99.8	435	2	US-09-968-415-6
4	2337	99.8	437	2	US-09-851-588-8
5	2297.5	98.1	432	2	US-10-012-331A-275
6	2297.5	98.1	432	2	US-10-015-389A-275
7	2297.5	98.1	432	2	US-10-006-768A-275
8	2297.5	98.1	432	2	US-10-015-671A-275
9	2297.5	98.1	432	2	US-10-015-933A-275
10	2297.5	98.1	432	2	US-10-011-833A-275
11	2297.5	98.1	432	2	US-10-006-041A-275
12	2297.5	98.1	432	2	US-10-012-064A-275
13	2186	93.4	406	2	US-09-656-002-2
14	2186	93.4	406	2	US-09-851-588-6
15	1266	54.1	292	2	US-09-607-745-9
16	777	33.2	418	2	US-10-177-661-6
17	698.5	29.8	477	2	US-10-177-661-2
18	697.5	29.8	562	2	US-09-879-792-12
19	683	29.2	446	2	US-10-177-661-4
20	677.5	28.9	492	2	US-09-885-166A-895
21	677.5	28.9	492	2	US-09-879-792-14
22	677.5	28.9	492	2	US-09-679-426-895
23	677.5	28.9	492	2	US-09-759-143-895
24	677.5	28.9	492	2	US-10-012-896-895
25	676.5	28.9	492	2	US-09-342-749-2
26	676.5	28.9	492	2	US-09-691-840-2
27	676.5	28.9	510	2	US-09-949-016-11074

28	675.5	28.8	393	2	US-09-759-143-934	Sequence 934, App
29	675.5	28.8	393	2	US-10-012-896-934	Sequence 934, App
30	675.5	28.8	492	2	US-09-759-143-932	Sequence 932, App
31	675.5	28.8	492	2	US-10-012-896-932	Sequence 932, App
32	660.5	28.2	521	2	US-09-949-016-11081	Sequence 11081, A
33	660.5	28.2	521	2	US-09-949-016-11082	Sequence 11082, A
34	660.5	28.2	521	2	US-09-949-016-11083	Sequence 11083, A
35	658	28.1	453	2	US-09-999-833A-69	Sequence 69, App1
36	658	28.1	453	2	US-10-020-445A-69	Sequence 69, App1
37	655.5	28.0	454	2	US-09-518-046-2	Sequence 2, App1
38	655.5	28.0	454	2	US-09-650-371-2	Sequence 2, App1
39	628.5	26.8	445	2	US-09-856-371A-8	Sequence 8, App1
40	588.5	25.1	417	2	US-09-820-002-4	Sequence 4, App1
41	584.5	25.0	452	2	US-09-949-016-7182	Sequence 7182, Ap
42	584	24.9	455	2	US-09-261-416-2	Sequence 2, App1
43	580	24.8	376	2	US-09-820-002-2	Sequence 2, App1
44	574	24.5	416	1	US-09-000-846-2	Sequence 2, App1
45	574	24.5	457	2	US-09-856-371A-10	Sequence 10, App1

ALIGNMENTS

RESULT 1									
US-09-607-745-2									
; Sequence 2, Application US/09607745									
; Patent No. 6750034									
; GENERAL INFORMATION:									
; APPLICANT: Dartow, Andrew L									
; APPLICANT: Qi, Jai-shen									
; APPLICANT: Andrade-Gordon, Patricia									
; TITLE OF INVENTION: DNA encoding human serine protease D-G									
; FILE REFERENCE: ORT-1273									
; CURRENT APPLICATION NUMBER: US/09/607,745									
; CURRENT FILING DATE: 2000-06-30									
; NUMBER OF SEQ ID NOS: 9									
; SOFTWARE: Patentin Ver. 2.1									
; SEQ ID NO 2									
; LENGTH: 435									
; TYPE: PRT									
; ORGANISM: Homo sapiens									
US-09-607-745-2									
Query Match									
Best Local Similarity 100.0%; Pred. No. 1.2e-230;									
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	MDPDSQPLNSLDVFKLRKP	1	MDPDSQPLNSLDVFKLRKP	1	MDPDSQPLNSLDVFKLRKP	1	MDPDSQPLNSLDVFKLRKP	1
DB	1	MDPDSQPLNSLDVFKLRKP	1	MDPDSQPLNSLDVFKLRKP	1	MDPDSQPLNSLDVFKLRKP	1	MDPDSQPLNSLDVFKLRKP	1
QY	61	LCGQPLHPIPRKOLCDGELDPLGDEDEHCYKSPPEGAVAVRISKORSTIQVDSATGN	61	LCGQPLHPIPRKOLCDGELDPLGDEDEHCYKSPPEGAVAVRISKORSTIQVDSATGN	61	LCGQPLHPIPRKOLCDGELDPLGDEDEHCYKSPPEGAVAVRISKORSTIQVDSATGN	61	LCGQPLHPIPRKOLCDGELDPLGDEDEHCYKSPPEGAVAVRISKORSTIQVDSATGN	61
DB	61	LCGQPLHPIPRKOLCDGELDPLGDEDEHCYKSPPEGAVAVRISKORSTIQVDSATGN	61	LCGQPLHPIPRKOLCDGELDPLGDEDEHCYKSPPEGAVAVRISKORSTIQVDSATGN	61	LCGQPLHPIPRKOLCDGELDPLGDEDEHCYKSPPEGAVAVRISKORSTIQVDSATGN	61	LCGQPLHPIPRKOLCDGELDPLGDEDEHCYKSPPEGAVAVRISKORSTIQVDSATGN	61
QY	121	WFSACFDNFETALAEACROWGYSKPTFAVEIGPDODLVEITENSQELMRNNSGP	121	WFSACFDNFETALAEACROWGYSKPTFAVEIGPDODLVEITENSQELMRNNSGP	121	WFSACFDNFETALAEACROWGYSKPTFAVEIGPDODLVEITENSQELMRNNSGP	121	WFSACFDNFETALAEACROWGYSKPTFAVEIGPDODLVEITENSQELMRNNSGP	121
DB	121	WFSACFDNFETALAEACROWGYSKPTFAVEIGPDODLVEITENSQELMRNNSGP	121	WFSACFDNFETALAEACROWGYSKPTFAVEIGPDODLVEITENSQELMRNNSGP	121	WFSACFDNFETALAEACROWGYSKPTFAVEIGPDODLVEITENSQELMRNNSGP	121	WFSACFDNFETALAEACROWGYSKPTFAVEIGPDODLVEITENSQELMRNNSGP	121
QY	181	CLSGLVSLHCLACGKSLKTRPVVCGEASVDSWPMOVSIOYDKQHYCGGSIIDPHWLT	181	CLSGLVSLHCLACGKSLKTRPVVCGEASVDSWPMOVSIOYDKQHYCGGSIIDPHWLT	181	CLSGLVSLHCLACGKSLKTRPVVCGEASVDSWPMOVSIOYDKQHYCGGSIIDPHWLT	181	CLSGLVSLHCLACGKSLKTRPVVCGEASVDSWPMOVSIOYDKQHYCGGSIIDPHWLT	181
DB	181	CLSGLVSLHCLACGKSLKTRPVVCGEASVDSWPMOVSIOYDKQHYCGGSIIDPHWLT	181	CLSGLVSLHCLACGKSLKTRPVVCGEASVDSWPMOVSIOYDKQHYCGGSIIDPHWLT	181	CLSGLVSLHCLACGKSLKTRPVVCGEASVDSWPMOVSIOYDKQHYCGGSIIDPHWLT	181	CLSGLVSLHCLACGKSLKTRPVVCGEASVDSWPMOVSIOYDKQHYCGGSIIDPHWLT	181
QY	241	AAHCRKHTDVFNNKVRAGSDKLSPSLIAVAKITIIIEPNMYPKNDIALMKLOPULTF	241	AAHCRKHTDVFNNKVRAGSDKLSPSLIAVAKITIIIEPNMYPKNDIALMKLOPULTF	241	AAHCRKHTDVFNNKVRAGSDKLSPSLIAVAKITIIIEPNMYPKNDIALMKLOPULTF	241	AAHCRKHTDVFNNKVRAGSDKLSPSLIAVAKITIIIEPNMYPKNDIALMKLOPULTF	241
DB	241	AAHCRKHTDVFNNKVRAGSDKLSPSLIAVAKITIIIEPNMYPKNDIALMKLOPULTF	241	AAHCRKHTDVFNNKVRAGSDKLSPSLIAVAKITIIIEPNMYPKNDIALMKLOPULTF	241	AAHCRKHTDVFNNKVRAGSDKLSPSLIAVAKITIIIEPNMYPKNDIALMKLOPULTF	241	AAHCRKHTDVFNNKVRAGSDKLSPSLIAVAKITIIIEPNMYPKNDIALMKLOPULTF	241
QY	301	SGTVRPICLPFDDELTPATPLWTIIGGFTKONGKMSDILQASVOVINDSTRCNADAY	301	SGTVRPICLPFDDELTPATPLWTIIGGFTKONGKMSDILQASVOVINDSTRCNADAY	301	SGTVRPICLPFDDELTPATPLWTIIGGFTKONGKMSDILQASVOVINDSTRCNADAY	301	SGTVRPICLPFDDELTPATPLWTIIGGFTKONGKMSDILQASVOVINDSTRCNADAY	301
DB	301	SGTVRPICLPFDDELTPATPLWTIIGGFTKONGKMSDILQASVOVINDSTRCNADAY	301	SGTVRPICLPFDDELTPATPLWTIIGGFTKONGKMSDILQASVOVINDSTRCNADAY	301	SGTVRPICLPFDDELTPATPLWTIIGGFTKONGKMSDILQASVOVINDSTRCNADAY	301	SGTVRPICLPFDDELTPATPLWTIIGGFTKONGKMSDILQASVOVINDSTRCNADAY	301
QY	361	QGVTEKMKMGITPEGVDTCQSGGPLMTYQSDQMHVGVIVSGYCGGPGSTPGVYTKV	361	QGVTEKMKMGITPEGVDTCQSGGPLMTYQSDQMHVGVIVSGYCGGPGSTPGVYTKV	361	QGVTEKMKMGITPEGVDTCQSGGPLMTYQSDQMHVGVIVSGYCGGPGSTPGVYTKV	361	QGVTEKMKMGITPEGVDTCQSGGPLMTYQSDQMHVGVIVSGYCGGPGSTPGVYTKV	361
DB	361	QGVTEKMKMGITPEGVDTCQSGGPLMTYQSDQMHVGVIVSGYCGGPGSTPGVYTKV	361	QGVTEKMKMGITPEGVDTCQSGGPLMTYQSDQMHVGVIVSGYCGGPGSTPGVYTKV	361	QGVTEKMKMGITPEGVDTCQSGGPLMTYQSDQMHVGVIVSGYCGGPGSTPGVYTKV	361	QGVTEKMKMGITPEGVDTCQSGGPLMTYQSDQMHVGVIVSGYCGGPGSTPGVYTKV	361

THIS PAGE BLANK (USPTO)

QY 402 ValSerTrpGlyTyrGlyCysGlyProSerThrProGlyValTyrThrIysValSer 421
Db 1484 GTTAGCTGGGGCTATGCTGCGGGGCCGAGCACCCCAAGATATACACCAAGTCTCA 1543
QY 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpIysAlaGluLeu 435
Db 1544 GCCTATCTCACTGATGATCTACAAATGCTGGAAGGCTGAGCTG 1585

Search completed: December 6, 2005, 20:22:41
Job time : 710 secs

PD 19-SEP-2002.
 XX
 PF 14-MAR-2002; 2002WO-US007826.
 XX
 PR 14-MAR-2001; 2001US-0276025P.
 PR 14-MAR-2001; 2001US-0276026P.
 PR 10-AUG-2001; 2001US-0311732P.
 PR 19-SEP-2001; 2001US-0323580P.
 PR 26-SEP-2001; 2001US-0324967P.
 PR 26-SEP-2001; 2001US-0325102P.
 PR 26-SEP-2001; 2001US-0325149P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 PI Monahan JE, Gannavarapu M, Hoersch S, Kamatkar S, Kovatis SG,
 PI Meyers RE, Morrisey MP, Qiantd PJ, Sen A, Vleby PO, Mills GB,
 PI Baer RC, Lu K, Schmandt RE, Zhao X, Glatt K;
 XX
 DR WPI: 2002-723277/78.
 DR P-PSDB; ABG96430.
 XX
 PT Assessing whether a patient is afflicted with ovarian cancer, useful in
 PT assessing the stage or progression of the disease, comprises comparing
 PT the expression level of a cancer marker in a sample from a patient and
 PT from a non cancer patient.
 XX
 PS Disclosure; Page 438; 481pp; English.
 XX
 CC The present invention relates to a new method for assessing whether a
 CC patient is afflicted with ovarian cancer. The method involves comparing
 CC the expression level of a marker in a patient sample and the normal level
 CC of expression of the marker in a control non-ovarian cancer sample, where
 CC the marker is selected from 363 cancer markers described in the
 CC specification. The method of the invention is useful in diagnosing or
 CC characterizing cancer, in detecting the presence of cancer as early as
 CC possible, and the recurrence of ovarian cancer. The method may also be of
 CC particular use with patients having an enhanced risk of developing
 CC ovarian cancer (e.g. patients having a familial history of ovarian
 CC cancer). The cancer markers may be used in the management and treatment
 CC of e.g. brain and central nervous system disorders (e.g. bacterial and
 CC viral meningitis, Alzheimer's disease or Parkinson's disease), brain
 CC disorders (e.g. cerebral oedema, hydrocephalus or brain herniations),
 CC inflammations (e.g. bacterial or viral meningitis or encephalitis),
 CC testicular disorders (e.g. nonbucculous granulomatous orchitis),
 CC connective tissue disorders, or heart disorders (e.g. ischaemic heart
 CC disease or atherosclerosis). The compositions and methods may also be
 CC used in assessing the histological type of neoplasm associated with
 CC ovarian cancer, monitoring the progression of ovarian cancer, determining
 CC whether ovarian cancer has metastasized or is likely to metastasize,
 CC selecting a composition for inhibiting ovarian cancer, assessing the
 CC ovarian carcinogenic potential of a compound, or inhibiting ovarian
 CC cancer or at risk of developing ovarian cancer. The present nucleic acid
 CC invention encodes one of the ovarian cancer markers described in the
 CC
 XX
 SQ Sequence 2307 BP; 555 A; 647 C; 614 G; 460 T; 0 U; 31 Other;
 XX
 Alignment Scores:
 Pred. No.: 6,95e-207 Length: 2307
 Score: 2337.00 Matches: 434
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 99.79% Indels: 0
 DB: 6 Gaps: 0
 US-10-803-530-2 (1-435) x ABS76529 (1-2307)
 QY 2 AAPPProAASPSeApgInProLeuAAnSerLeuAAspValIySPProLeuAArgIySPProAArg 21
 Db 284 GATCCGAGACAGTCATCAACTCTGGAACAGGCTCGATGTCACAAACCCCTGCGCAAAACCCCGT 343
 QY 22 TlePrometGluThrPheArgIySValGlyIleProIleIleIleIleAleuSerLeu 41

Db	344	ATCCCAATGGAGACCTTCAGAAAGGTGGGGATCCCATCATCATAGACACTAGAGCTGT	403
OY	42	AlaserlletlellevalValValleuilelysValilleuabpylserTyrPheleu	61
Db	404	GCGAGTATCATCATTTGGTGTGTCTCATCAAGGTGATTTCTGTAAATACACTTCTCTC	463
OY	62	Cys61IyGlnProleuWhlspheLleProArgly62IleuCyasbpgly61IleuAsbCys	81
Db	464	TGCGGGACGCTCTCCCACTTCATCTCCAGAGGAAGCGCTGTGTGACGGAGACTGACTGT	523
OY	82	ProleuGIyGluAsbpglyGluLuhIscYsValIySerPheProGIyGlyProAlaValAla	101
Db	524	CCCTTGGGGGAGGACGAGAGACCTGTGTCAAGAGCTTCCCCGAAGGGCTTCGACTGGCA	583
OY	102	ValArgleuSerlyasbArgSerThrIleuGlnValIleuabSeralathrGIyAsnTyr	121
Db	584	GTCCGGCTCTCCAGAGACCAGTATCCACATCGAGGTGTGACTCGGCCACAGGAACTGG	643
OY	122	PheSeralAcYsPheasbAsnPhethrGIuAlaleuAlagIuThrAlaCyAsArgGlnMet	141
Db	644	TTCTGTGCTGTTCGACAACTTCACAGAACCTCTGCTGTGAAGACGCTGTAGGCAAGTGTG	703
OY	142	GIyTyrSerSerlybProthrPheArgAlaValGlnIeglyProAsbpglAsbPleuAsp	161
Db	704	GCTACACAGACAAACCACTTTCAGACTGTGAGATTTGGCCCAAGACAGAGATCTGCAT	763
OY	162	ValValGlnIethrGIuAsbSerGIuIleuArgMetArgAsbSerSerGIyProCys	181
Db	764	GTTGTGTGAATACAGAAACAGCAGGAGCTTGGATCGGGAACTCAAGTGGGGCCGTGT	823
OY	182	LeuSerGIySerIleuValSerIleuHiscYsleuAlaCy61IylySerIleuIyThrPro	201
Db	824	CTCTCAGGCTCCCTGGCTCTCCCTCAGCTGTCTCTGCTGTGGGAAGAGCTTGAAGCCCC	883
OY	202	ArgValValGIyGIyGIuGlnAlaSerValasbSerTyrProTpgIInvalISerIleGln	221
Db	884	CGTGTGGTGGGTGGAGAGAGCCCTCTGTGGATTTCTGGCCTTGGACAGTACGATCCAG	943
OY	222	TyrAsbpylsglnHlsvalCy61GIySerIleuAsbProHlSTyrValIleuThrAla	241
Db	944	TACGACAAACAGCAGCTGTGTGAGAGGAGCATCTCGAACCCCACTGGGTCTTCACGGCA	1003
OY	242	AlahiscYsPheArglyHlsIthrAsbValPheasbTpyIysValArgAlaGIySerAsp	261
Db	1004	GCCCACTGCTTCAGGAAACATACCGATGTGTTCACATCGAAGAGTGGCGGACAGCTCAGAC	1063
OY	262	LysIleuGIySerPheProSerIleuAlaValAlaIySellellelleGluPheAsnPro	281
Db	1064	AAACTGGGACACTTCCATCCCTGGCTGTGGCCAAAGATCATCATTTGAATTCAAACCC	1123
OY	282	MetTyrProLybAsbAsnAspIleAlaleuMetIySleuGlnPheProLeuThrPheSer	301
Db	1124	ATGTACCCCAAGACATATGACATGGCCCTTCATGAAGCTGCAGTTCCACTCACTTTCTCA	1183
OY	302	GIyThrValArgProIleCyAsleuProPhePheAsbpglyIuIleuThrProAlaThrPro	321
Db	1184	GGCAACAGTCAGGCCCATCTGTCTGGCCCTTTTGTATGAGGAGCTCAGTCCAGCCACCCCA	1244
OY	322	LeuTrrIlelleGIyTrrGIyPheThrIySglnAsngIyGIyIySmetSerAspIleu	341
Db	1244	CTCTGGATCATATGGATGGGCTTTACGAAAGACATGAGAGGAAGATCTTCGACATCTACTG	1303
OY	342	IleuGlnIasbServalGlnValIleAsbSerThrArgCyasbAlaAsbAspAlaTyrGln	361
Db	1304	CTGCAGGGGCTCAGTCCAGGTATTTGACAGCACGCTGCATTCAGAGAGATGCCATCCAG	1364
OY	362	GIyIuValIThrGIyIySmetMetCySAlagIyIleProGIuGIyGIyValAsbThrCys	381
Db	1364	GGGGAAGTCAACGAGAAAGATGATGTGTGCAGGACATCCGGAAGGGGGGTGTGCACCTGC	1423
OY	382	GlnGIyAsbSerGIyGIyProIleuMetIyGlnSerAsbpglnTrrHlsValValGIyIle	401
Db	1424	CAGGTGTGACAGTGTGGGCCCTGTATGTACAAATCTGACAGTGGCAATGTGTGTGGCACTC	1483

XX Ashkenazi A, Goddard A, Gurney A, Polakis P, Smith V, Wood WI;
 PI Wu T, Zhang Z;
 XX WPI; 2005-418022/42.
 DR P-PSDB; AEA54849.
 XX
 PT New polynucleotides encoding tumor-associated antigenic target (TAT)
 PT polypeptides, useful e.g., in gene therapy, in diagnosis of cancer, and
 PT in the development of cancer therapeutics.
 XX
 PS Claim 2; SEQ ID NO 2; 137bp; English.
 XX
 CC The invention relates to novel tumor associated antigenic target (TAT)
 CC polypeptides and nucleic acid molecules encoding such polypeptides.
 CC Nucleic acid molecules of the invention are useful as hybridization
 CC probes in chromosome and gene mapping, for producing TAT polypeptides, in
 CC gene therapy for treating TAT expressing tumors and in the diagnosis of
 CC such tumors. They are also useful for producing transgenic animals for
 CC screening and developing therapeutically useful reagents, in tissue
 CC typing and for producing probes for PCR, Northern analysis, Southern
 CC analysis and Western analysis. The present sequence is the human TAT502
 CC cDNA.
 XX
 SQ Sequence 2104 BP; 499 A; 604 C; 577 G; 424 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 6,15e-207 Length: 2104
 Score: 2337.00 Matches: 434
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 99.79% Indels: 0
 DB: 14 Gaps: 0
 US-10-803-530-2 (1-435) x AEA54844 (1-2104)
 QY 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValIysProLeuAlaGlySerProAla 21
 Db 242 GATCTGACGATGATCAACCTCTGAAACAGCTCGATGTCAAAACCCCTGCGCAAAACCCCGT 301
 QY 22 IleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSerLeu 41
 Db 302 ATCCCAATGAGACCTTCAGAAAGGTGGGATCCCATCATCATCATCATCATCATCATCATCAT 361
 QY 42 AlaserIleIleIleValIleValIleValIleValIleValIleValIleValIleValIleVal 61
 Db 362 GCGAGTATCATCATTTGGTGTGCTTCATCAAGGTATCTGGATTAATAACTACTTCCG 421
 QY 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys 81
 Db 422 TGGGGGAGGCTCTCCACCTTCATCCCGAGAGACACTGTGTGACGAGAGCTGGACTGT 481
 QY 82 ProLeuGlyGlnAspGluGluHisCysValIysSerPheProGluGlyProAlaValAla 101
 Db 482 CCTTGGGGAGGACGAGGACGACTGTGTCAAGAGCTTCCCGAAGGGCTGTGCACTGCA 541
 QY 102 ValArgLeuSerLysAspArgSerThrLeuGlnValIleAspSerAlaThrGlyAsnTrp 121
 Db 542 GTCCGCTCTCCAGAGACCGATTCACATGCAAGGTGCTGACTCCGCCACAGGGAACTGG 601
 QY 122 PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
 Db 602 TTCTGTGCTGTTTGCACAACTTCACAGAGCTCTCGGTGACAGCCCTGTAGGCGAATG 661
 QY 142 GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
 Db 662 GGTCTAGAGGAAACCCCACTTTCAGAGCTGTGGAGATTGGCCGACCAAGACTCTGAT 721
 QY 162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
 Db 722 GTTGTGAATCAGAGAAACGCGAGAGCTTGCATGCGAACTCAAGGGGCGCTGT 781
 QY 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro 201

Db 782 CTCACAGGCTCCCTGGTCTCCCTGCACCTGTCTTGCTGTGGAGAGAGCTTGAAGACCCCC 841
 QY 202 ArgValValGlyIleGluGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
 Db 842 CGTGTGGTGGTGGGAGAGAGGCTCTGTGATTTCTTGCCCTTGGCAGGTGACATCCAG 901
 QY 222 TyrAspLysGlnHisValCysGlyLysSerIleLeuAspProHisTrpValLeuThrAla 241
 Db 902 TACGACAAACAGACAGCTCTGTGAGAGACATCTTGACCCCACTGGAGTCTTCAACGCA 961
 QY 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
 Db 962 GCCCACTGTCTTCAGAAACATACCATGATGTTCACCTGAAGGTGGCGGAGCTCAAGC 1021
 QY 262 LysLeuGlySerPheProSerIleuAlaValAlaLysIleIleIleIleGluPheAsnPro 281
 Db 1022 AACTGGGCGAGCTTCCATCCCTGGCTGTGGCCAAAGATCATCATCATCATCATCATCAT 1081
 QY 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
 Db 1082 ATGTACCCCAAGACATGACATGACATGACATGACATGACATGACATGACATGACAT 1141
 QY 302 GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro 321
 Db 1142 GGCACAGTCAGGCCCATCTGTCTGCTCTTCTTGTATGAGAGAGCTCACTCAGGCAACCCA 1201
 QY 322 LeuTrpIleIleGlyTyrPglPheThrLysGlnAsnGlyIleLysMetSerAspIleLeu 341
 Db 1202 CTCTGGATCATTTGATGGGCTTTTACAGACGAGATGAGAGAGAGATGCTGTGACATACAG 1261
 QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGln 361
 Db 1262 CTGAGGGGTCAAGCTCAGATCATGACAGACACGGGTGCAATGACATGATGTGTCCAG 1321
 QY 362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyIleValAspThrCys 381
 Db 1322 GGGGAAATCACCGAAGAGATGATGTGACAGGATCCCGAAGGGGTGTGACACCTGC 1381
 QY 382 GlnGlyAspSerGlyIleProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle 401
 Db 1382 CAGGTGACAGTGTGGGCGCCCTGATGTCATCATGACATGATGATGTGTGTGGCAATC 1441
 QY 402 ValSerTrpGlyTyrGlyCysGlyIleProSerThrProGlyValTyrThrLysValSer 421
 Db 1442 GTTAGCTGGGGCTATGAGCTGGCGGGGCCCGAGACCCCGAGATATACCAAGGTCTCA 1501
 QY 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
 Db 1502 GCCATCTCAACTGATCTTCAATGTCTGGAAGGCTGAGCTG 1543
 RESULT 15
 ID ABS76529 standard; cDNA; 2307 BP.
 XX
 AC ABS76529;
 XX
 DT 11-DEC-2002 (first entry)
 XX
 DE cDNA encoding human ovarian cancer marker OV86.
 XX
 KW Human; ovarian cancer; marker; cancer; familial history; brain disorder;
 KW central nervous system disorder; bacterial meningitis; viral meningitis;
 KW Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;
 KW brain herniation; inflammation; encephalitis; testicular disorder;
 KW nonbacterial granulomatous orchitis; connective tissue disorder;
 KW heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;
 KW histological type; carcinogenic; ovarian cancer marker; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200271928-A2.

Db	302	ATCCCATGAGACCTTCAGAAAGTGGGATCCCATCATCATAGACATCTGAGCTCG	361	Db	1382	CAGGCTACAGTGTGGGCCCTTGATGTACCAATCTGACACAGTGCATGTGGGCATC	1441
Qy	42	Alaserllellellevalvalleuileysvalilleuaspblystrtryrphleu	61	Qy	402	Valsertrpglytryrglycysglylproserthrproglyvaltrythrlyvalser	421
Db	362	GCGAGTATCATCATGTGTGTCTCATAGAGTATTTCTGAATAATCTACTTCCCTC	421	Db	1442	GTTAGCTGGGGCTRTAGCTCTGGGGGCCGAGACATCCCGAGGTATACACCAAGTCTCA	1501
Qy	62	CysglylnproleuwhispehilleproarglysglnleucysaspglyluleuaspCys	81	Qy	422	AlatyrleuanttrpillertryaenvaltrplysallagluLeu	435
Db	422	TGCGGGAGACCTCTCCACTTCATCCCGAGAGACAGCTGTGACCGAGAGCTGACCTGT	481	Db	1502	GCCATCTCAATCGATCTACATATGTCTGGAAGGCTGAGCTG	1543
Qy	82	ProleuglylulaspglulwhisCysvallyseserPheproglulglyproalavalala	101	RESULT 13			
Db	482	CCCTTGGGGGAGGACGAGAGCACTGTGTCAAGAGCTTCCCGAAGGCGCTCAGTGGCA	541	ID	ADM04403	standard; cDNA, 2104 BP.	
Qy	102	ValArgleuserlysaaspargserthrleuglnvalleuaspseralanthrglyasnrtp	121	XX	ADM04403;		
Db	542	GTCGGCTCTCCAAAGCCGATCCACATCGAGGTCTGATCTGGCCACAGGGAACTGG	601	XX	07-APR-2005	(first entry)	
Qy	122	PheSerAlaCysPheaspbnphethrgulalaleuallgluthralaCysArgglnmet	141	XX			
Db	602	TTCTGTCCCTGTTTTCACAACTTCACAGAGCTCTGCTGAGACAGCTGTAGGCATG	661	DE	Human Ovarian specific nucleic acid, Ovr115.		
Qy	142	GlyTyrSerSerlysaProthrPheargalaValglulleglProaspglulaspLeuasp	161	KW	Ovarian specific nucleic acid; OSNA; ss; gene; chromosome-11; cytostatic;		
Db	662	GGCTACAGCAGCAACCCACTTTCAGAGCTGTGAGATTGGCCACAGACAGGATCTGAT	721	XX	neoplasm; ovary tumor; Ovr115; SNP; single nucleotide polymorphism.		
Qy	162	ValValglullethrGlulaspserglululeuargmetArgasnseserGlyProCys	181	OS	Homo sapiens.		
Db	722	GTTGTGTAATACAGAAACAGCCAGAGGCTTCCCATGCGGAACCTCAAGTGGCCCTGT	781	FH	Key	Location/Qualifiers	
Qy	182	LeuserglyserleuValserleuwhisCysleuallCysglylyseserleuYthrPro	201	FT	CDS	233..1546	
Db	782	CTCTCAGGCTCCCTGTCTCCCTGCACTGTCTGTGTGGAAAGCTGAGACCCCTC	841	FT		/*tag= a	
Qy	202	ArgValvalglylglulglulalaserValaspsertrpProtpglnvalserlleugln	221	FT	variation	/product= "Ovr115"	
Db	842	CGTGTGGTGGTGGGAGAGAGGCTCTGTGATCTTGTGGCTTGGCAGTCCAGATCCAG	901	FT	variation	/tag= b	
Qy	222	TyrAspLysGlnhisvalCysglylyserlleuaspProhiStrpValleuThrAla	241	FT	variation	/tag= c	
Db	902	TACACAAACAGACGCTGTGAGAGACATCTGTGACCCCACTGGGTCTCAGGCA	961	FT	variation	/standard_name= "Single nucleotide polymorphism"	
Qy	242	AlahisCysPhearglyshisethrAspValPheasntrpIystrpValargllyserAsp	261	FT	variation	/*tag= e	
Db	962	GCCCACTGCTTCAAGAAACATACCGATGTGTTCACATCGAAGGTGGGGAGGCTCAGAC	1021	FT	variation	/standard_name= "Single nucleotide polymorphism"	
Qy	262	LysleuGlySerPheProserleuAlaValAlalysllellellelleglupheAsnPro	281	FT	variation	/*tag= f	
Db	1022	AAACTGGGAGCTTCCCATCTGGCTGTGGCCAGATCATCATTTGAAATTCACACCC	1081	FT	variation	/standard_name= "Single nucleotide polymorphism"	
Qy	282	MetTyrProLysAspAsnAspillealaleuMetLysleuGlnPheProleuThrPheSer	301	FT	variation	/*tag= g	
Db	1082	ATGTACCCCAAAAGCAATGACATGCCCTCATTAACCTGACATCCACTCTTCTCA	1141	FT	variation	/standard_name= "Single nucleotide polymorphism"	
Qy	302	GlyThrValArgProilleCysleuProPhePheaspglulLeuThrProalThrPro	321	FT	variation	/*tag= g	
Db	1142	GGCAGAGTCAGGCCATCTGTCTGCCCTTCTTGAAGAGACTCATCTCAACCCCA	1201	FT	variation	/standard_name= "Single nucleotide polymorphism"	
Qy	322	LeuTrpillellellytrpIyPheThrlysglnaenGlylyLysMetSeraspilleu	341	FT	variation	/standard_name= "Single nucleotide polymorphism"	
Db	1202	CTCTGGATCATTTGGATGGGCTTTTACGAAGCAATGAGGGGAATGTCTGACACTCTG	1261	FT	variation	/standard_name= "Single nucleotide polymorphism"	
Qy	342	LeuGlnAlaserValGlnValilleaspserThrArgCysAsnAlaAspAspAlaTryGln	361	FT	variation	/standard_name= "Single nucleotide polymorphism"	
Db	1262	CTGCAAGGCTCAAGTCAATTCACACACACAGCTGCAATGACAGATGCTTACACAG	1321	FT	variation	/standard_name= "Single nucleotide polymorphism"	
Qy	362	GlyGlyValThrGlyLysMetMetCysalaglylleProglulgllyValAspThrCys	381	FT	variation	/standard_name= "Single nucleotide polymorphism"	
Db	1322	GGGGAAGTCACCGAAGATGATGTGTGAGCAATCCCGAAGGGGGTGTGACACTGCTC	1381	FT	variation	/standard_name= "Single nucleotide polymorphism"	
Qy	382	GlnGlyAspSerGlyGlyProleuMetLysglinsarpglntrpHisvalValgllyle	401	FT	variation	/standard_name= "Single nucleotide polymorphism"	

The invention relates to an isolated nucleic acid molecule (ovarian specific nucleic acid). Also included are determining the presence of a ovarian specific nucleic acid (OSNA) in a sample, a vector comprising an OSNA, a host cell comprising the vector, producing a polypeptide encoded

```
Db      224 GATCTACAGTATCAACCTCTGAACAGCTCGATGTCAAACCCCTCGCGAAACCCCGT 283
Qy      22  ILePrometGluThrPheArgLysValGlyLeProIleIleIleIleValLeuLeuSerIleu 41
Db      284 ATCCCATGAGACCTTCAGAAAGTGGGGATCCCATCATCATAGACACTGAGCTG 343
Qy      42  ALaSerIleIleIleValValLeuIleLysValIleLeuAspLysIleValIlePheLeu 61
Db      344 GCGAGTATCATCATGTGTGTGTCTCATCAAGGTGATTTCTGTAATAATCTACTTCTTC 403
Qy      62  Cy6eLysGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys 81
Db      404 TGCGGAGAGCTCTCCACTTCATCCAGAGAGAGAGCTGTGTGACGAGAGACTGAGCTGT 463
Qy      82  ProLeuGlyGluAspGlyGluIleIleCysValLysSerPheProGlyGluProAlaValAla 101
Db      464 CCTTGGGGGAGAGAGAGAGAGACTGTGTCAAGAGCTTCCCGAAAGGGCTGCGAGTGGCA 523
Qy      102 ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp 121
Db      524 GTCCGCTCTCCAGAGAGCCATCCACTGACAGGTGCTGAGCTGCGCCACAGGGAACCTGG 583
Qy      122 PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
Db      584 TTCTGTCTCTTTGACAACTTCACAAAGCTCTCGGTGAGACAGCTGTAGGACAGATG 643
Qy      142 GLYTYrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlyAspLeuAsp 161
Db      644 GGCTACAGACGAAACCCACTTTCAGAGCTGTGAGATTGGCCAGACCAAGACTGTGAT 703
Qy      162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
Db      704 GTTGTGAATACAGAAACAGCCAGAGAGCTTGGCACTCGGAAGCTCAAGTGGGCTGT 763
Qy      182 LeuSerGlySerLeuValSerIleuHisCysLeuAlaCysGlyLysSerLeuLysThrPro 201
Db      764 CTCACAGGCTCCCTGTCTCCCTGCACTGTCTTGGGAGAGAGCTGAGAGAGAGAGAGAG 823
Qy      202 ArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
Db      824 CGTGTGTGGGTGGGAGAGAGAGCTCTGTGATTTCTGAGCTTGGCAGGTGACGATCCAG 883
Qy      222 TYrAspLysGlnHisValCysGlyLysSerIleLeuAspProHisTrpValLeuThrAla 241
Db      884 TACGACAAACAGCAGCTCTGTGAGAGAGAGCATCTCGAGAGAGAGAGAGAGAGAGAGAG 943
Qy      242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
Db      944 GCCCATCTCTCAGGAAACATACGATGTTCCTCACTGGAAGGTGCGGGAGAGCTCAGAC 1003
Qy      262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleIleGluPheAsnPro 281
Db      1004 AAACGGGAGAGCTTCCATCTCTGCTGTGGGCAAGATCATCATGTAATTCAACCCC 1063
Qy      282 MetTYrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db      1064 ATGTACCCCAAGAAATGACATCGCCCTCATGAGAGCTGACATCTCCACTCTTCTCA 1123
Qy      302 GlyThrValArgProIleCysLeuProPhePheAspGlyGluLeuThrProAlaThrPro 321
Db      1124 GGCAACAGTACAGGCCATCTGTCTGCCCTTCTGTATGAGAGAGACTCACTCCAGCCACCA 1183
Qy      322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAnGlyLysLysMetSerAspIleLeu 341
Db      1184 CTCTGATCATGATGAGTGGGCTTTTACGAAGCAGAAATGGAGGGAAGATGTCTACATACG 1243
Qy      342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTYrGln 361
Db      1244 CTGCGAGGCTCAGTCCAGGTCAATTGACACCAACGGTGCAGAGACCATCGTACCG 1303
Qy      362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyValAlaAspThrCys 381
```

```
Db      1304 GGGGAAGTCACCGAGAAGATGATGTGTGACAGCATCCCGAAGGGGTGTGACACCTGC 1363
Qy      382 GlnGlyAspSerGlyGlyProLeuMetTYrGlnSerAspGlnTrpHisValIleGlyIle 401
Db      1364 CAGGGTACAGTGTGTGGCCCTGATGTACCAATCTGACAGTGGATGTGTGGGATC 1423
Qy      402 ValSerTrpGlyTYrGlyCysGlyGlyProSerThrProGlyValTYrThrLysValSer 421
Db      1424 GTTAGCTGGGGCTATGGCTCGGGGGGCCGAGACCCCAAGAGTATACACCAAGTCTCA 1483
Qy      422 AlaTYrLeuAsnTrpIleTYrAsnValTrpLysAlaGluLeu 435
Db      1484 GCCATCTCAACTGAGATCTCAATGTCTGGAAGGCTGAGCTG 1525

RESULT 12
ADN04863 standard; cDNA; 2104 BP.
AC ADN04863;
AC 01-JUL-2004 (first entry)
DE Antipsoriatic cDNA sequence #646.
KM ds; gene; antipsoriatic; gene therapy; psoriasis; diagnosis.
OS Homo sapiens.
PN MO2004028479-A2.
PD 08-APR-2004.
XX 25-SEP-2003; 2003WO-US030907.
XX 25-SEP-2002; 2002US-0414006P.
PR (GERTH ) GENENTECH INC.
PA Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI,
PI Wu TD;
PI WPI; 2004-305105/28.
DR P-PSDB; ADN04864.
XX New PRO nucleic acid or polypeptide, useful for preparing a
PT pharmaceutical composition for diagnosing or treating psoriasis in a
PT mammal.
XX Claim 1; SEQ ID NO 1257; 3069bp; English.
XX The invention relates to novel polynucleotide and polypeptides for
CC treating psoriasis or a sequence having at least 80% identity to the
CC above sequences. The nucleic acid is useful for preparing a composition
CC for diagnosing or treating psoriasis in a mammal. This sequence
CC corresponds to one of the polynucleotides of the invention.
XX
SQ Sequence 2104 BP; 499 A; 604 C; 577 G; 424 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 6,15e-207 Length: 2104
Score: 2337.00 Matches: 434
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.79% Indels: 0
DB: 12 Gaps: 0

US-10-803-530-2 (1-435) x ADN04863 (1-2104)
Qy      2  AspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuAlaArgLysProArg 21
Db      242 GATCTGACAGTATCAACCTCTGAACAGCTCGATGTCAAACCCCTCGCGAAACCCCGT 301
Qy      22  ILePrometGluThrPheArgLysValGlyLeProIleIleIleIleValLeuLeuSerIleu 41
```

Db 310 GTCCGCTCTCCAGAGCCGATCCACACTGCTGAGTCTGGACTCGGCGACAGAGAACTGG 369
 QY 122 PheSerAlaCysPheAspSerPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
 Db 370 TTCTCTGCGCTGTTTCCACAATTCCACAGAGCTCTGCTGAGACAGCCGTGAGCAGATG 429
 QY 142 GlyTyrSerSerIleProThrPheArgAlaValGluLeuGlyProAspGlnAspLeuAsp 161
 Db 430 GGCTACAGCAGCAAAACCCACTTTCAGAGCTGTGGAGATTGGCCACAGACAGATCTGGAT 489
 QY 162 ValValGluIleThrGluAsnSerGlnIleLeuArgMetArgAsnSerSerGlyProCys 181
 Db 490 GTTGTGTAATACAGAAACAGCAGAGAGCTTCCATGCGGAATCTCAAGTGGGCGCTGT 549
 QY 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro 201
 Db 550 CTCTCAGAGCTCCCTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 609
 QY 202 ArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
 Db 610 CGTGTGTGGTGGGAGAGAGGCTCTGTGATCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 669
 QY 222 TyrAspLysGlnHisValCysGlyLysSerIleLeuAspProHisTrpValLeuThrAla 241
 Db 670 TACGACAAACAGCAGTCTGTGAGAGGAGCATCTGAGACCCCACTGGGTCTTACGCGCA 729
 QY 242 AlaHisCysPheArgLysHisThrAspValPheSerTrpLysValAlaGlyLysSerAsp 261
 Db 730 GCCCACTGCTTCCAGGAAACATACCGATGTGTTCAACTGAGAGGAGGAGGAGGAGGAGG 789
 QY 262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleIleGluPheAsnPro 281
 Db 790 AAACCTGGGAGCTTCCCATCCCTGCTGTGGCCAAAGTATCATCATTAATTCACACCC 849
 QY 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
 Db 850 ATGTACCCCAAGCAATGACATGCCCCATGATGACATGATCCCATCTTCTCA 909
 QY 302 GlyThrValArgProIleCysLeuProPhePheAsnGlnIleuThrProAlaThrPro 321
 Db 910 GGACACTCAGGCCCATCTGTGCTCTTCTTGTGAGAGGAGTCACTCAGACCAACCCCA 969
 QY 322 LeuTrpIleIleGlyTyrGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu 341
 Db 970 CTCTGATCATGTGATGGGCTTTACGAGACAAATGAGAGGAGATGTTCTGACATACCTG 1029
 QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGln 361
 Db 1030 CTGCAAGCGTCAGTCCAGGTCAATTCAGACACCGGTGCAATGACAGATCGTACCCAG 1089
 QY 362 GlyLysValThrGlnLysMetMetCysAlaGlyIleProGlnGlyGlyValAspThrCys 381
 Db 1090 GGGGAAGTACCCAGAAAGATGATGTGTGAGGAGCATCCCGAAAGGGGTGTGACACCTGC 1149
 QY 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle 401
 Db 1150 CAGGGTGAAGTGTGGGCGCCCGTGAATGCAATTCAGCACTGGCAGTGGTGGGCTC 1209
 QY 402 ValSerTrpGlyTyrGlyCysGlyGlyProSerThrTrpGlyValTyrThrLysValSer 421
 Db 1210 GTTAGCTGGGCTATGGCTGGCGGGGCGCGAGCACCCAGAGATTTACACCAAGGTCTCA 1269
 QY 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
 Db 1270 GCTTATCTCACTGATCTACATGTCTGAAAGCTGAGCTG 1311
 RESULT 11
 AA172976
 ID AA172976 standard, cDNA, 2081 BP.
 XX
 AC AA172976;
 XX
 DT 21-AUG-2002 (first entry)

XX CXA8 preferred cDNA.
 DE Gene: colorectal cancer; CGA7, CJA8; modulating protein; screening;
 XX drug candidate; vaccine; ss.
 KW Homo sapiens.
 OS
 XX
 PH Key Location/Qualifiers
 FT CDS 215..1528
 FT /tag= a
 XX /product= "CJA8"
 PN US2002042067-A1.
 PD 11-APR-2002.
 PF 08-MAY-2001; 2001US-00851588.
 PR 17-AUG-2000; 2000US-00642252.
 PR 06-SEP-2000; 2000US-00656002.
 XX (MACK/) MACK D.
 PA (GISH/) GISH K C.
 PA (WILS/) WILSON K E.
 XX
 PI Mack D, Gish KC, Wilson KE;
 DR MPI. 2002-453647/48.
 DR P-PSDB; AAG79359.
 XX
 PT Screening drug candidates for treating colorectal cancer, comprises
 PT determining the effect of the candidate on the expression profile gene of
 PT CGA7 or CJA8.
 PS Claim 69, Fig 7; 40pp; English.
 XX
 CC The sequences given in AA172973-76 encode the colorectal cancer proteins,
 CC CGA7 and CJA8. The CGA7 and CJA8 proteins are colorectal cancer
 CC modulating proteins and have been mapped to chromosomes 2 (CGA7) and 11
 CC (CJA8). These sequences may be used in the method of the invention for
 CC screening drug candidates. The method comprises adding a drug candidate
 CC to a cell that expresses an expression profile gene encoding CGA7, CJA8
 CC or fragments and determining the effect of the drug candidate on the
 CC expression of the expression profile gene. The new methods are used to
 CC screen bioactive agents for the ability to bind to or modulate the
 CC activity of CGA7 or CJA8 and evaluate the effect of a candidate
 CC colorectal cancer drug. An antibody to CGA7 or CJA8 can inhibit the
 CC activity of CGA7 or CJA8, respectively, and is used to screen for an
 CC agent that can interfere with the binding of CGA7 or CJA8 to the
 CC antibody. The antibody can be used to treat colorectal cancer. The
 CC colorectal cancer tissue, where the therapeutic group is a cytotoxic
 CC agent or a radiolabeled. Antisense molecules are used to inhibit
 CC colorectal cancer in a cell. Nucleic acid segments encoding CGA7 or CJA8
 CC are used in a biochip. CGA7, CJA8 or a nucleic acid encoding it are used
 CC to elicit an immune response. CGA7 or CJA8 is used to determine the
 CC prognosis of an individual with colorectal cancer. Nucleic acid encoding
 CC CGA7 or CJA8 can be used in vaccines
 XX
 SQ Sequence 2081 BP; 484 A; 597 C; 576 G; 424 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 6,06-207
 Score: 2337.00 Length: 2081
 Percent Similarity: 100.00% Matches: 434
 Best Local Similarity: 100.00% Conservative: 0
 Query Match: 99.79% Mismatches: 0
 DB: 6 Indels: 0
 Gaps: 0
 US-10-803-530-2 (1-435) x AA172976 (1-2081)
 QY 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuAlaGlyLysProArg 21

Oy	222	TyrAspIysGlnHisValCysGlyGlySerIleLeuAspProHisIleThrValLeuThrAla	241
Db	670	TACGACAAACACGACGCTGTGTGAGAGGACATCTCGACCCCCCATGTGGTCTCTCACGGCA	729
Oy	242	AlaHisCysPheArgLysHisIleThrAspValPheAsnTPlysValArgAlaGlySerAsp	261
Db	730	GCCCACTCTCTTACGAAACAATACGATGTGTTCACTGGAAAGTGTGGGGCAGGCTCAAC	789
Oy	262	LysLeuGlySerPheProSerLeuAlaValAlaValIleIleIleIleGluPheAsnPro	281
Db	790	AAACTGGGCGAGCTTCCCATCCCTGGCTGTGGCCAGACATCATCATGAATTCAACCCC	849
Oy	282	MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer	301
Db	850	ATGTACCCCAAGAAAGAAATGACATCCCTCTCATGAAGCTGACAGTTCCACTCATTTTCTCA	909
Oy	302	GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro	321
Db	910	GGCACAGTCAGAGCCCATCTGTCTGTCCCTTTTGTATGAGAGCTACATCCAGCCACCCA	969
Oy	322	LeuThrIleIleGlyTyrGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu	341
Db	970	CTCTGATCATGTGATGGGGCTTTTACGAAGCAGATGAGGGAGATGTCTGACATACGTG	1029
Oy	342	LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGln	361
Db	1030	CTGCAGGGGCTCAGTCCAGGTCACTTATACAGCACACGGTGTCAATGACAGAGATGGCTTACAG	1089
Oy	362	GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys	381
Db	1090	GGGAGAGTCACCGAAGAATGATGTGTGACGACATCCCGAAAGGGGGTGTGACACCTGCG	1149
Oy	382	GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnThrPheIleValIleGlyIle	401
Db	1150	CAGGGTGACAGTGGTGGGCCCTGTATGATACCAATCTGACACAGTGGCATGTGGTGGGCATC	1209
Oy	402	ValSerTPGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSer	421
Db	1210	GTTAGCTGGGCTAATGGCTGCGGGGGCCCGACACCCAGAGATATACCAACCAAGGTCTCA	1269
Oy	422	AlaTyrLeuAsnThrIleTyrAsnValTPlysAlaGluLeu	435
Db	1270	GCCTATCTCACTGATCTACATGTCTTGAAAGCGCTGAGCTG	1311
RESULT 10			
ADU25739			
ID	ADU25739	standard; cDNA; 1314 BP.	
AC	ADU25739;		
XX			
DT	27-JAN-2005	(first entry)	
XX			
DE		cDNA encoding human TMPRSS4 protease.	
XX			
KM		Human; transmembrane serine protease 4; TMPRSS4; cardiovascular disorder;	
KM		endocrinological disease; gastroenterological disease;	
KM		urological disorder; metabolic disease; cancer; respiratory disease;	
KM		dermatological disease; cardiac; antiarteriosclerotic; antidiabetic;	
KM		hypertensive; antiinflammatory; laxative; nephrotoxic; nootropic;	
KM		osteopathic; neuroprotective; antiparkinsonian; cyostatic; gene; ss.	
XX			
OS		Homo sapiens.	
XX			
Key		Location/Qualifiers	
FT	CDS	1..1314	
FT		/*tag= a	
FT		/product= "TMPRSS4 protease"	
XX			
PN	WO2004097034 -A2.		
XX			
PD	11-NOV-2004.		
XX			

```

PF 20-APR-2004; 2004MO-EP004164.
XX
XX 02-MAY-2003; 2003EP-00010004.
XX
XX (FARB ) BAYER HEALTHCARE AG.
XX
XX Gold S, Bruggemeier U, Geerts A, Polej S;
PI WPI; 2004-804771/79.
XX
XX P-PSDB; ADU25740.
XX
XX Screening for therapeutic agents, useful for treating cancer,
PT cardiovascular, endocrinological, urological, metabolic, respiratory, or
PT dermatological diseases, comprises contacting a test compound with a
PT transmembrane serine protease 4.
XX
XX Disclosure; SEQ ID NO 1; 134pp; English.
XX
XX The invention relates to a method of screening for therapeutic agents
XX useful in the treatment of human disease. The method comprises contacting
XX a test compound with a transmembrane serine protease 4 (TMPRSS4)
XX polypeptide or polynucleotide, or determining the activity of a TMPRSS4
XX polypeptide at a certain concentration of a test compound or in the
XX absence of the test compound. The method is useful for diagnosing,
XX treating, and preventing cardiovascular disorders (e.g. myocardial
XX infarction, arrhythmias, hypertensive or peripheral vascular diseases, or
XX atherosclerosis), endocrinological diseases (e.g. diabetes mellitus,
XX Zollinger-Ellison syndrome, Addison's disease, Cushing's syndrome,
XX hyperaldosteronism, or thyroiditis), gastroenterological diseases (e.g.
XX gastritis, chronic inflammatory disease, Crohn's disease, diarrhea, or
XX constipation), urological disorders (e.g. renal disease,
XX glomerulopathies, nephritis, or obstructive uropathies), metabolic
XX diseases (e.g. obesity, Lesch-Nyhan syndrome, osteoporosis, Alzheimer's
XX disease, Parkinson's disease, or Fanconi's syndrome), cancer, respiratory
XX diseases (e.g. asthma or chronic obstructive pulmonary disease) and
XX dermatological diseases (e.g. psoriasis, ichthyosis, keratosis pilaris,
XX psoriasis), pityriasis rosea, or dermatitis). This sequence encodes human
XX TMPRSS4 protease.
XX
XX
XX Sequence 1314 BP; 297 A; 369 C; 373 G; 275 T; 0 U; 0 Other;
XX
XX
XX Alignment Scores:
XX Pred. No.: 3,3e-207 Length: 1314
XX Score: 2237,00 Matches: 434
XX Percent Similarity: 100,00% Conservative: 0
XX Best Local Similarity: 100,00% Mismatches: 0
XX Query Match: 99,79% Indels: 0
XX DB: 13 Gaps: 0
XX
XX US-10-803-530-2 (1-435) x ADU25739 (1-1314)
XX
XX 2 ASPPQAPSPSASAPGGLNPROLEUMENSERLEUASPVALYSPROLEUNRGYSPROARG 21
XX 10 GATCTGTGACAGATCAACCTTGAAACACCTCGAAGTCAAAACCTGTGGCAAAACCCCGT 69
XX
XX 22 ILEPROMETGLNTHPHEARGLYSVALGIYILEPROLIELEILEALEULEUSERLEU 41
XX 70 ATCCCAATGGAACCTTCAGAAAGTGGGGAATCCCATATCATATACACTACTGAGGCTG 129
XX
XX 42 ALASERILEILEILEVALVALLEULEILEYVALILEULEUSPLYTYTYRPHLEU 61
XX 130 GCGAATATCATCATTTGTGGTTCTCCATCAAGAGTGATTTCTGGATAAATCTACTCTC 189
XX
XX 62 CYSGLYGLNPROLEUNHISPHLEILEPROADGLYGLINLEUCYASPLGYLULEUAPCY 81
XX 190 TGGGGGACGCGCTCCACTTCATCCGAGAGAACACTGTGTGACGAGAGCTGACCTGT 249
XX
XX 82 PROLEUNGLYGLUASPGUGLUHISCYVALYSSERPHAPROGLUGLYPROALAVALALA 101
XX 250 CCGTTGGGGAGAGACGAGAGACACTGTGTCAAGAGCTTCCCGAAGGCGCTGACGTGCA 309
XX
XX 102 VALAAGLEUSERLYASPARGSERTHRILEUGINVALLEUASPSEVALTHRGLYASNTTP 121

```


Db 1090 GGGGAAGTACCGAGAAATGATGTGTGCGAGCATCCCGAAGGGGTGTGGACATCTGC 1149
 QY 382 GNGIYASpSerGlyGlyProLeuMetTyrGlnSerAspGlnThrHisValValGlyIle 401
 Db 1150 CAGGGTGACAGTGTGTGGCCCTGATGTCCATCTGACAGTGTGTGTGGGATC 1209
 QY 402 ValSerTyrGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrIleValSer 421
 Db 1210 GTTAGCTGGGGCTATGTGCTGCGGGGGCCGAGCACCCGAGAGTATACCAAGGTCTCA 1269
 QY 422 AATyrLeuAsnTrpIleTyrAsnValTrrpValIleu 435
 Db 1270 GCTATCTCACTGATCTGATCTGAAATGTGGAAGGCTGAGCTG 1311
 RESULT 9
 ADN39460
 ID ADN39460 standard; cDNA; 1314 BP.
 AC ADN39460;
 XX
 DT 17-JUN-2004 (first entry)
 XX
 DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:A60.
 XX
 KW Human; differential expression; cancer; angiogenic disorder;
 KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
 KW inflammatory disease; autoimmune disease;
 KW rectal neovascularisation syndrome; scarring; uterine fibroid;
 KW detection; diagnosis; prognosis; drug screening; drug targeting;
 KW wound healing; contraception; cytosstatic; cardiant; immunomodulatory;
 KW vulnery; gene therapy; vaccine; gene; ss.
 XX
 OS Homo sapiens.
 XX
 EN WO2003042661-A2.
 PD 22-MAY-2003.
 XX
 PF 13-NOV-2002; 2002MO-US036810.
 XX
 PR 13-NOV-2001; 2001US-0350666P.
 PR 21-NOV-2001; 2001US-033246P.
 PR 29-NOV-2001; 2001US-034393P.
 PR 03-DEC-2001; 2001US-0335394P.
 PR 14-DEC-2001; 2001US-0340376P.
 PR 08-JAN-2002; 2002US-0347349P.
 PR 10-JAN-2002; 2002US-0347211P.
 PR 08-FEB-2002; 2002US-0355250P.
 PR 13-FEB-2002; 2002US-0356714P.
 PR 20-FEB-2002; 2002US-0359077P.
 PR 29-MAR-2002; 2002US-036809P.
 PR 04-APR-2002; 2002US-0370110P.
 PR 12-APR-2002; 2002US-0372246P.
 PR 05-JUN-2002; 2002US-038614P.
 PR 16-JUL-2002; 2002US-0396839P.
 PR 22-JUL-2002; 2002US-039775P.
 PR 22-JUL-2002; 2002US-0397845P.
 PR 09-SEP-2002; 2002US-0409450P.
 XX
 PA (E05B-) EOS BIOTECHNOLOGY INC.
 XX
 PI Afar D, Aziz N, Ginsburg WM, Gish KC, Glyme R, Hevazi PA,
 PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
 XX
 DR MPI; 2003-468649/44.
 XX
 DX P-PSDB; ADN39461.
 XX
 PT Determining the presence or absence of a pathological cell in a patient,
 PT useful for diagnosing, prognosing or treating cancer, comprises detecting
 PT a nucleic acid in a biological sample.
 XX
 PS Claim 8; SEQ ID NO A60; 1385bp; English.
 XX

CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
 CC whose expression is upregulated or downregulated in specific cancers or
 CC other diseases such as angiogenic or fibrotic disorders, and to methods
 CC of determining the presence or absence of a pathological cell in a
 CC patient by detecting a nucleic acid at least 80% identical to those of
 CC the invention or by detecting a polypeptide of the invention. The
 CC invention also relates to expression vectors and host cells comprising a
 CC nucleic acid of the invention; antibodies which specifically bind a
 CC polypeptide of the invention; use of such antibodies for drug targeting;
 CC and methods of screening for modulators of activity or expression of the
 CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
 CC antibodies and methods are useful for diagnosing, prognosing and treating
 CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
 CC atherosclerosis, inflammatory diseases, autoimmune diseases, rectal
 CC neovascularisation syndromes, scarring and uterine fibroids. They may
 CC also be useful in wound healing and in contraception. The present
 CC sequence represents a nucleic acid sequence of the invention.
 CC
 SQ Sequence 1314 BP; 297 A; 369 C; 373 G; 275 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 3,3e-207 Length: 1314
 Score: 2337.00 Matches: 434
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 99.79% Indels: 0
 DB: 11 Gaps: 0
 US-10-803-530-2 (1-435) x ADN39460 (1-1314)
 QY 2 AspProAspSerAspGlnProLeuAsnSerLeuAspVallyProLeuArgProArg 21
 Db 10 GATCTGACAGTGTATCAACCTTGACAGCTTCATGTCAACCCCTGCGCAACCCCGT 69
 QY 22 IleProMetGluThrPheArgLysValGlyIleProIleIleIleIleLeuLeuSerLeu 41
 Db 70 ATCCCATGGAGACCTTCAGAAAGTGGAGATCCCATCATCATATGACATCTGACCTG 129
 QY 42 AlaSerIleIleIleValValIleuIleLysValIleLeuAspLysTyrTyrPheLeu 61
 Db 130 GCGAGTATCATATTGGTGTGTCCTCATCAAGTGTGATGATTAATATCTTCCCTC 189
 QY 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyLeuAspCys 81
 Db 190 TCCGGAGAGCTCTCCACTTCATCCGAGAACCACTGTGTGACCGAGAGGTGACTGT 249
 QY 82 ProLeuGlyLeuAspGluGluHisCysValLysSerPheProGluGlyProAlaValAla 101
 Db 250 CCTTGGGGAGAGACAGAGAGCACTGTGTCAAGAGCTTCCCGAAGGGCTTGCAGTGGCA 309
 QY 102 ValArgLeuSerLysAspArgSerThrLeuGlnValIleuAspSerAlaThrGlyAsnTrp 121
 Db 310 GTCGCTCTCCCAAGAGCCGATCCACATGAGGTCTGTGACTCGGCCACAGGAACTGG 369
 QY 122 PheSerAlaCysPheAspAsnPhetrrgualaleuAlaGluThrAlaCysArgIleMet 141
 Db 370 TTCTCTGCTCTTTTCACAACTTCACAGAGCTCTCTGACAGACCGCTGTGCGCAGATG 429
 QY 142 GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeu 161
 Db 430 GGCTACAGACAGCAACCACTTTCAGAGCTGTGAGATTTGGCCAGACCAAGATCTGGAT 489
 QY 162 ValValGluIleThrGluAsnSerGlnLeuArgMetArgAsnSerSerGlyProCys 181
 Db 490 GTTGTGAATATCACAAACAGCAGGAGCTTCCGATCGGAATCAAGTGGGCTGT 549
 QY 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro 201
 Db 550 CTCTACAGCTCTCTGCTCTCTGACATGTGTGCTGTGGGAAAGCTTAAAGCCCCC 609
 QY 202 ArgValValGlyGlyGluGluLaseValAspSerTrpProTrpGlnValSerIleGln 221
 Db 610 CGTGTGTGGTGGGAGAGGCTCTGTGGAATCTTGGCTTGGCAGGTACGATCCAG 669

XX Homo sapiens.
OS
XX MO2003042661-A2.
PN
XX
PD 22-MAY-2003.
XX
PF 13-NOV-2002; 2002WO-US0366810.
XX
PR 13-NOV-2001; 2001US-0350666P.
PR 21-NOV-2001; 2001US-0332464P.
PR 29-NOV-2001; 2001US-0334393P.
PR 03-DEC-2001; 2001US-0335394P.
PR 14-DEC-2001; 2001US-0340376P.
PR 06-JAN-2002; 2002US-0347211P.
PR 10-JAN-2002; 2002US-0347349P.
PR 08-FEB-2002; 2002US-0355250P.
PR 13-FEB-2002; 2002US-0356714P.
PR 20-FEB-2002; 2002US-0359077P.
PR 29-MAR-2002; 2002US-036809P.
PR 04-APR-2002; 2002US-0370110P.
PR 12-APR-2002; 2002US-0372246P.
PR 05-JUN-2002; 2002US-0386614P.
PR 16-JUL-2002; 2002US-0396839P.
PR 22-JUL-2002; 2002US-0397775P.
PR 22-JUL-2002; 2002US-0397845P.
PR 09-SEP-2002; 2002US-0409450P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
PI Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnick A;
XX WPI; 2003-468649/44.
DR P-PSDB; ADN38816.
XX
PT Determining the presence or absence of a pathological cell in a patient,
PT useful for diagnosing, prognosing or treating cancer, comprises detecting
PT a nucleic acid in a biological sample.
XX
XX
XX Claim 8; SEQ ID NO 133; 1385bp; English.
XX
XX The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC whose expression is upregulated or downregulated in specific cancers or
CC other diseases such as angiogenic or fibrotic disorders, and to methods
CC of determining the presence or absence of a pathological cell in a
CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC invention also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularisation syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a nucleic acid sequence of the invention.
XX
XX Sequence 1314 BP; 297 A; 369 C; 373 G; 275 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 3.3e-207 Length: 1314
Score: 2337.00 Matches: 434
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.79% Indels: 0
DB: 11 Gaps: 0
US-10-803-530-2 (1-435) x ADN38815 (1-1314)
Qy 2 AspProAspSerAspGlnProLeuAnSerLeuAspValysProLeuArgIysProArg 21

Db 10 GATCCTGACAGTGTATCAACCTCTGAACAGCCTCGATGCAAAACCCCTGCGCAACCCCGT 69
Qy 22 ILeProMetGluThrPheArgIysValGlyIleProIleIleIleIleIleLeuSerLeu 41
Db 70 ATCCCATGGAGACCTTCAGAAAGTGGGATGCCCATCATCATGACATGACATGAGCCTG 129
Qy 42 ALaSerIleIleIleValIleValIleuIleIysValIleLeuAspIysTyrTyrPheLeu 61
Db 130 GCGAGTATCATCATTTGGTGTCTCTCATCANGTGTATTCGAGTAATACTACTTCTTC 189
Qy 62 CysGlyGlnProLeuHisPheIleProArgIysGlnLeuCysAspGlyGluLeuAspCys 81
Db 190 TGGCGGAGAGCCTCTCCACTTCATCCAGAGAAAGCAGCTGTGTGACGGAGAGCTGACTGT 249
Qy 82 ProLeuGlyGluAspGlyGluHisCysValIysSerPheProGluGlyProAlaValAla 101
Db 250 CCTTTGGGGGAGGACGAGGAGCAGTGTCTCAAGAGCTTCCCGAAGGGCCCTGACGTGCA 309
Qy 102 ValArgLeuSerIysAspArgSerThrLeuGlnValIleuAspSerAlaThrGlyAsnTrp 121
Db 310 GTCCGCTCTTCAGAGACCATTCACACTGAGAGTCTGGACTGGCCACAGGAACTGG 369
Qy 122 PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
Db 370 TTCTCGCTGTGTTTCGAACTTCACAGAGCTCTCGTGAGACAGCCTGTAGGACAGATG 429
Qy 142 GlyTyrSerSerIysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
Db 430 GGCTACAGACGAAACCCACTTTCAGAGCTGTGAGATGGCCAGACCCAGACTCTGAT 489
Qy 162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
Db 490 GTGTGTAATCAGAAAGAAACAGCAGAGAGCTGTGACTCCGAATCAAGTGGCCCTGT 549
Qy 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyIysSerLeuIysThrPro 201
Db 550 CTCACAGCTCTCCGCTCTCTCCCTGCACTGTCTCTGCGTGGAAAGAGCTGAAGCCCCC 609
Qy 202 ArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
Db 610 CGTGTGGGTGGGTGGGAGAGCCTCTGTGATTTCTGGCCTTGGCAGGTGAGCATTCAG 669
Qy 222 TyrAspIysGlnHisValCysGlyIysSerIleLeuAspProHisTrpValLeuThrAla 241
Db 670 TACACAAACGACAGCTGTGGAGGAGCATCTGAGCCCCACTGGGTCTCAAGGCA 729
Qy 242 AlaHisCysPheArgIysHisThrAspValPheAsnTrpIysValArgAlaGlySerAsp 261
Db 730 GCCACCTCTTCAGAAACATACCGATGTCTCACTGAAAGTGGCGGAGGCTCAGAC 789
Qy 262 LysLeuGlySerPheProSerLeuAlaValAlaIysIleIleIleGluPheAsnPro 281
Db 790 AAATCGGGCAGCTTCCCATCTCGGTGGCCAAAGATCATCATCATTAATTCAACCCC 849
Qy 282 MetTyrProIysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db 850 ATGTACCCCAAGCAATGACATGCTCTATGAAGCTGACGTTCCACTCATCTTCTCA 909
Qy 302 GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro 321
Db 910 GGCACAGTCAGGCCCATGTGCTGCCCTCTTTATAGAGAGCTCACTCCAGCCACCCCA 969
Qy 322 LeuTrpIleIleGlyTyrGlyPheThrIysGlnAsnGlyIysIysMetSerAspIleLeu 341
Db 970 CTCTGGATCATTTGATGGGGCTTTTACGAGCAGATGAGGAAAGATCTTGACATCTG 1029
Qy 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGln 361
Db 1030 CTGCAAGGCTGATGCTCAGCTCATTCAGCACACCGTGCATATGACAGATGCGTACAG 1089
Qy 362 GlyGluValThrGluIysMetMetCysAlaGlyIleProGluGlyValAspThrCys 381

PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnick A;
 XX WPI; 2003-468649/44.
 DR P-PsDB; ADN39473.
 XX
 PT Determining the presence or absence of a pathological cell in a patient,
 PT useful for diagnosing, prognosing or treating cancer, comprises detecting
 PT a nucleic acid in a biological sample.
 XX
 PS Claim 8; SEQ ID NO A72; 1385bp; English.

XX The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
 CC whose expression is upregulated or downregulated in specific cancers or
 CC other diseases such as angiogenic or fibrotic disorders, and to methods
 CC of determining the presence or absence of a pathological cell in a
 CC patient by detecting a nucleic acid at least 80% identical to those of
 CC the invention or by detecting a polypeptide of the invention. The
 CC invention also relates to expression vectors and host cells comprising a
 CC nucleic acid of the invention; antibodies which specifically bind a
 CC polypeptide of the invention; use of such antibodies for drug targeting;
 CC and methods of screening for modulators of activity or expression of the
 CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
 CC antibodies and methods are useful for diagnosing, prognosing and treating
 CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
 CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
 CC neovascularisation syndromes, scarring and uterine fibroids. They may
 CC also be useful in wound healing and in contraception. The present
 CC sequence represents a nucleic acid sequence of the invention.
 XX

SQ Sequence 1314 BP; 297 A; 369 C; 373 G; 275 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 3,3e-207 Length: 1314
 Score: 2337.00 Matches: 434
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 99.79% Indels: 0
 DB: Gaps: 0

US-10-803-530-2 (1-435) x ADN39472 (1-1314)

QY 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValIysProLeuArgIysProArg 21
 Db 10 GATCTCGACAGATATCAACCTCTGAACAGCCTCGATGTAACCCCTGGGCAAAACCCCT 69
 QY 22 IleProMetGlnThrPheArgIysValIleProIleIleIleAlaLeuLeuSerIleu 41
 Db 70 ATCCCATGAGACCTTCAGAAAGTGGGATCCCATCATCATCTACTGAGCCTG 129
 QY 42 AlaSerIleIleIleValValIleuIleIleValIleLeuAspIysThrIlePheIleu 61
 Db 130 GCGAGTATCATCTATGGTGTGCTCCATCATCAAGTATCTGATTAATTAATCTTCC 189
 QY 62 CysGlyGlnProLeuHisPheIleProArgIysGlnLeuCysAspGlyGluLeuAspCys 81
 Db 190 TGGGGGACACCTCTCCACTTCATCCGAGGAAGACACTGTGCGAGGAGCTGACTGT 249
 QY 82 ProLeuGlyGluAspGluGlnHisCysValIysSerPheProGluGlyProAlaValAla 101
 Db 250 CCTTGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTCCCGAAGGGGCTGCACTGCA 309
 QY 102 ValArgLeuSerIysAspArgSerThrLeuGlnValIleuAspSerAlaThrGlyAsnThr 121
 Db 310 GTCCGCTCTCCAAAGGACCATCCACATCGAGGTGCTGAGCTCGGCCACAGGAACTGG 369
 QY 122 PheSerAlaCysPheAspAsnThrThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
 Db 370 TTCTCTGCTGTTTGACAAACCTTCAAGAGCTCTCGCTGAGACAGCCTGTAGAGAGATG 429
 QY 142 GlyTyrSerSerIysProThrPheArgIleValIleGluIleGlyProAspGlnAspLeuAsp 161
 Db 430 GCGTACAGAGCAACCCACTTTCAGAGCTGTGAGATTGGCCCGACAGACAGATCTGAT 489

QY 162 ValValGluIleThrGlnAsnSerGlnIleuArgMetArgAsnSerSerGlyProCys 181
 Db 490 GTTGTGAATTCACAAACACAGGAGCTTCGATGCGAATCAAGTGGGCTGT 549
 QY 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyIysSerLeuThrPro 201
 Db 550 CTCTCAGGCTCCCTGGTCTCCCTGACATGCTGTGCTGTGGGAAGCTGAAGACCCC 609
 QY 202 ArgValValGlyGluGlnGluAlaSerValAspSerThrProThrGlnValSerIleGln 221
 Db 610 CGTGTGTGGTGGGAGGAGGCTCTGTGATTTTGGCTTGGCAGGTGAGATTCAG 669
 QY 222 TyrAspIysGlnHisValCysGlyIysSerIleLeuAspProHisThrValLeuThrAla 241
 Db 670 TACGACAAACACACAGTCTGTGTGAGGAGACATCTCGAACCCCACTGGGATCTCAAGCA 729
 QY 242 AlaHisCysPheArgIysHisThrAspValPheAsnThrIysValArgAlaGlySerAsp 261
 Db 730 GCCCACTGCTCAGAAACATACCATGATGTTCACCTGAAAGTGGCGGCGCTCAGAC 789
 QY 262 LysLeuGlySerPheProSerIleuAlaValAlaIleIleIleIleGluPheAsnPro 281
 Db 790 AAATCGGCGAGCTTCCATCCCTGCTGTGGCCAAAGATCATCATTAATTAACCCC 849
 QY 282 MetTyrProIysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
 Db 850 ATGTACCCCAAGACATATGATGATGCTCCATGAAAGCTGCACTTCCACTCTCTTCA 909
 QY 302 GlyThrValArgProIleCysLeuProPhePheAspGlnIleuThrProAlaThrPro 321
 Db 910 GGCACAGTCAGGCCCATCTGTCTGCTCCCTTCTTGTATGAGGAGCTCATCCAGCACCCA 969
 QY 322 LeuThrIleIleIleGlyTyrGlyPheThrLysGlnAsnGlyIysMetSerAspIleu 341
 Db 970 CTCTGAGATCATTTGATGAGGCTTTTACGAAAGATGAGGAGGAGATGTGTGACATCTG 1029
 QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGln 361
 Db 1030 CTGCAAGGCTCAGTCCAGGTGATTCATGACGACACGGTGCAATGACAGATGGTACAG 1089
 QY 362 GlyGluValThrGlyLysMetMetCysAlaGlyIleProGluGlyIysValAspThrCys 381
 Db 1090 GGGGAAGTCACCGAAGATGATGTGACGAGCATCCCGAAGGGGTGGACACCTGC 1149
 QY 382 GlnIysAspSerGlyGlyProLeuMetTyrGlnSerAspGlnThrPheValIleGlyIle 401
 Db 1150 CAGGTGACAGTGTGGGCTTGTATGATTCACATTCGACCAAGTGTGTGGGCATC 1209
 QY 402 ValSerThrGlyTyrGlyCysGlyIysProSerThrProGlyValTyrThrLysValSer 421
 Db 1210 GTTAGCTGGGGCTATGGCTCGGGGGCCGAGCACCCAGAGATATACCAAGATCTCA 1269
 QY 422 AlaTyrLeuAsnThrIleTyrAsnValThrLysAlaGluLeu 435
 Db 1270 GCTTATCTCAATGAGATTAACAATGTGTGAAGGCTGAGCTG 1311

RESULT 8
 ADN38815
 ID ADN38815 standard; cDNA; 1314 BP.
 XX
 AC ADN38815;
 XX
 DT 17-JUN-2004 (first entry)
 XX
 DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:133.
 XX
 KW Human; differential expression; cancer; angiogenic disorder;
 KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
 KW inflammatory disease; autoimmune disease;
 KW retinal neovascularisation syndrome; scarring; uterine fibroid;
 KW detection; diagnosis; prognosis; drug screening; drug targeting;
 KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
 KW vulnerary; gene therapy; vaccine; gene; ss.

Alignment Scores:

Pred. No.: 3,36-207 Length: 1314
 Score: 2337.00 Matches: 434
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 99.79% Indels: 0
 DB: 11 Gaps: 0

US-10-803-530-2 (1-435) x ADN39661 (1-1314)

QY 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValIysProLeuArgIysProArg 21
 Db 10 GATCTGACAGTATCAACCTCTGAACGCTTCATGCAAAACCCCTGGCAAAACCCGT 69
 QY 22 IleProMetGluThrPheArgIysValIleProIleIleIleAlaLeuLeuSerLeu 41
 Db 70 ATCCCCATGAGACCTTCAGAAAGGTGGGAGATCCCATCATATGACACTAGAGCTG 129
 QY 42 AlaSerIleIleIleValIleValIleLeuIleIysValIleLeuAspIysTyrThrPheLeu 61
 Db 130 GCGAGTATCATATTGGTGTCTCTCATCAAGGTGATTCGATAAATGACTACTCTC 189
 QY 62 CysGlyGlnProLeuHisPheIleProArgIysGlnLeuCysAspGlyIleLeuAspCys 81
 Db 190 TCGGGGAGCTCTCCACTTTCATCCCGAGAAAGAGCTGTGTGACGGAGAGCTGACTGT 249
 QY 82 ProLeuGlyGluAspGluGlnIleCysValIysSerPheProGluGlyProAlaValAla 101
 Db 250 CCTTTGGGGGAGAGAGAGAGACAGTGTCAAGAGCTTCCCGAAAGGGCTGGAGTGGCA 309
 QY 102 ValArgLeuSerIysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp 121
 Db 310 GTCCGCTCTCCAGAGCAGATCCACACTGAGGTGTGACTCGGCCACAGGGAACTCG 369
 QY 122 PheSerAlaCysPheAspAsnPheThrGlnAlaLeuAlaGluThrAlaCysArgGlnMet 141
 Db 370 TTCTGTGCTGTTCGACAACTTCACAGAGCTCTGCTGAGACAGCCGTGAGGAGATG 429
 QY 142 GlyTyrSerSerIysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
 Db 430 GGCTACACAGCAAAACCACTTCAGAGCTGTGGAGATTTGGGCCGACAGCATGTGAT 489
 QY 162 ValValGluIleThrGluAsnSerGlnIleuArgMetArgAsnSerSerGlyProCys 181
 Db 490 GTTGTGAATTCACAGAAACAGCAGAGAGCTTCGATGCGGAACTCAAGTGGGCTGT 549
 QY 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyIysSerLeuIleThrPro 201
 Db 550 CTCTGAGGCTCTCTGCTCTCCCTGCACTGTCTTGTGCTGTGGGAAAGAGCTTGAAAGCCCC 609
 QY 202 ArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
 Db 610 CGTGTGGTGGTGGGAGAGAGGCTCTGTGATTTCTTGCCCTTGGCAGGTCAGCATCCAG 669
 QY 222 TyrAspIysGlnHisValCysGlyIysSerIleLeuAspProHisIleTrpValLeuThrAla 241
 Db 670 TACGCAAAACAGCAGTCTGTGGAGGAGCATCCCGAACCCCACTGGGTCTCTCAGGGA 729
 QY 242 AlaHisCysPheArgIysHisIleThrAspValPheAsnTrpIysValArgAlaGlySerAsp 261
 Db 730 GCCCACTGCTTCAGAAACATACGATGTGTTCAACTGGAAAGGTCCGGGACAGGCTCAGAC 789
 QY 262 LysLeuGlySerPheProSerIleuAlaValIleIleIleIleGlnPheAsnPro 281
 Db 790 AAACGTGGGAGCTTCCCATCTGCTGTGGCCAAAGATATATATGAAATTCAAACCC 849
 QY 282 MetTyrProIysAspAsnAspIleAlaLeuMetIysLeuGlnPheProLeuThrPheSer 301
 Db 850 ATGTACCCCAAGACATAGCATCGCCCTCATGAAGCTGCAATTCCCATCTCTTCTCA 909
 QY 302 GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro 321
 Db 910 GGCACAGTCAGGCCCATGTGTGCTGCTTCTTGTATGAGAGACTCATCTCCAGCCACCCCA 969

QY 322 LeuThrIleIleIleGlyTrpGlyPheThrIysGlnAsnGlyIysIysMetSerAspIleLeu 341
 Db 970 CTCTGATCATTTGGATGGGGCTTTACAGAGCAGATGAGAGGAAAGATCTCGACATCTG 1029
 QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGln 361
 Db 1030 CTCGAGGCGTCAGTCCAGGTTCATTCAGACACAGCGTGCAATGACAGATGCGTACAG 1089
 QY 362 GlyGluValThrGluIysMetMetCysAlaGlyIleProGluGlyIysValAspThrCys 381
 Db 1090 GGGGAAATGACCCGAGAGATGATGTGACAGGCTATCCCGAAAGGGGTGTGGACCTGC 1149
 QY 382 GlnGlyAspSerGlyIysProLeuMetTyrGlnSerAspGlnTrpHisValIleGlyIle 401
 Db 1150 CAGGTACAGTGGTGGGCCCCCTGATATCCAACTCAACAGTGCAATGTGTGGCACTC 1209
 QY 402 ValSerTrpGlyTyrGlyCysGlyIysProSerThrProGlyValTyrThrIysValSer 421
 Db 1210 GTTAGCTGGGCTATGGCTGTGGGGGCCCGAGCACCCACAGAGTATACCAAGTCTCA 1269
 QY 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpIysAlaGluLeu 435
 Db 1270 GCTATCTCAACTGATCTACATGTCTGGAAGGCTGAGCTG 1311

RESULT 7
 ADN39472
 ID ADN39472 standard; cDNA; 1314 BP.
 XX AC ADN39472;
 XX DT 17-JUN-2004 (first entry)
 DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:472.
 KW Human; differential expression; cancer; angiogenic disorder;
 KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
 KW inflammatory disease; autoimmune disease;
 KW retinal neovascularization syndrome; scarring; uterine fibroid;
 KW detection; diagnosis; prognosis; drug screening; drug targeting;
 KW wound healing; contraception; cytosolic; cardiant; immunomodulatory;
 KW vulnerey; gene therapy; vaccine; gene; ss.
 OS Homo sapiens.
 XX PN WO2003042661-A2.
 XX PD 22-MAY-2003.
 XX PF 13-NOV-2002; 2002WO-US036810.
 XX PR 13-NOV-2001; 2001US-0350666P.
 XX PR 21-NOV-2001; 2001US-0332464P.
 XX PR 29-NOV-2001; 2001US-0334393P.
 XX PR 03-DEC-2001; 2001US-0335394P.
 XX PR 14-DEC-2001; 2001US-0340376P.
 XX PR 08-JAN-2002; 2002US-0347211P.
 XX PR 10-JAN-2002; 2002US-0347349P.
 XX PR 08-FEB-2002; 2002US-0355250P.
 XX PR 13-FEB-2002; 2002US-0356714P.
 XX PR 20-FEB-2002; 2002US-0359077P.
 XX PR 29-MAR-2002; 2002US-0368809P.
 XX PR 04-APR-2002; 2002US-0370110P.
 XX PR 12-APR-2002; 2002US-0372246P.
 XX PR 05-JUN-2002; 2002US-0386614P.
 XX PR 16-JUL-2002; 2002US-0396839P.
 XX PR 22-JUL-2002; 2002US-0397775P.
 XX PR 22-JUL-2002; 2002US-0397845P.
 XX PR 09-SEP-2002; 2002US-0409450P.
 PA (EOSB-) EOS BIOTECHNOLOGY INC.
 XX Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;

```

Db      310 GTCCGCTTCCAGAGCCGATCCACACGAGGCTGAGTCCGACCACAGGAACTGG 369
QY      122 PheSerAlaCySphEspAspPheThrGluAlaLeuAlaGluThrAlaCyArgGlnMet 141
Db      370 TTCTGCGCTGTGTGGACAACCTTCACAGAACCTTCCTGCGTAGAGACCTGTAGGCAATG 429
QY      142 GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
Db      430 GGCTACAGCAGCAAAACCTTCAGAGCTGTGAGATTGGCCAGACCAAGATCTGGAT 489
QY      162 ValValGluIleThrGluAsnSerGlnGluLeuAlaGluMetArgAsnSerSerGlyProCys 181
Db      490 GTTGTGAAATCACAGAAACAGCCAGAGCTTCCATGCGGAACTCAAGTGGGCCCTGT 549
QY      182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro 201
Db      550 CTCTCAGGCTCCCTGGTCTCCCTGACATGTCTGCTGTGGAAAGAGCTGAAAGACCC 609
QY      202 ArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
Db      610 CGTGTGGTGGGTGGGAGAGAGCCCTCTGTGATTCTTGCCCTTGCGACAGTCAGATCAG 669
QY      222 TyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAla 241
Db      670 TACACAAACAGCAGCTCTGTGAGAGGAGCATCTGAGACCCCTGGTCTTCACGGCA 729
QY      242 AlaHisCysPheArgLysHisThrAspAlaPheAsnTrpLysValAlaGluArgLysAsp 261
Db      730 GCCACGCTTCAAGAAACATACCCGATGTGTTCAACTGAAAGGTCGGGAGGCTCAGAC 789
QY      262 LysLeuGlySerPheProSerSerLeuAlaValAlaLysIleIleIleGluPheAsnPro 281
Db      790 AAATCGGAGCTTCCATCCCTGGCTGTGGCCAGATCATCATTAATTAATCAACCCC 849
QY      282 MetCysProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db      850 ATGTAACCCCAAGCAATGACATGCCCCATGAGCTGACATGCCACTTCTCTCA 909
QY      302 GlyThrValArgProIleCysLeuProPhePheAspGlnGluLeuThrProAlaThrPro 321
Db      910 GGACAGTCAGGCCCATCTCTGCTCTTGTGATGAGAGCTCATCCAGCACCCCA 969
QY      322 LeuTrpIleIleGlyTyrGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu 341
Db      970 CTCCTGATCATGTGAGGCTTTTACGAGCAGAAATGAGAGGAGATGTCTGACATACG 1029
QY      342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGln 361
Db      1030 CTGCAAGCGCTCAGTCCAGTCCATTTGACAGACACGCTGCAATGACAGATCGTACCG 1089
QY      362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys 381
Db      1090 GGGGAAGTCACCGAAAGATGATGTGTGAGAGGATCCGAAAGGGGTGTGACACCTGC 1149
QY      382 GlnGlyAspSerGlyGlyProLeuMetLysGlnSerAspGlnTrpHisValValGlyIle 401
Db      1150 CAGGGTGAAGTGTGGTGGCCCTGATGTACAAATCTGACCAATGGCATGTGTGGGATC 1209
QY      402 ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSer 421
Db      1210 GTTAGCTGGGCTATGCTGCGGGGCGCCGAGCACCCGAGGATATACCAAGATCTCA 1269
QY      422 AlaTyrLeuAsnTrpIleTyrAsnValTyrLysValIleGluLeu 435
Db      1270 GCCTATCTCACTGATCTCAATGATCTGGAAGGCTGAGCTG 1311

```

```

DT      17-JUN-2004 (first entry)
XX      XX
DB      Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO: C33.
XX      XX
KW      Human, differential expression; cancer; angiogenic disorder;
KW      fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW      inflammatory disease; autoimmune disease;
KW      retinal neovascularization syndrome; scarring; uterine fibroid;
KW      detection; diagnosis; prognosis; drug screening; drug targeting;
KW      wound healing; contraception; cytostatic; cardiac; immunomodulatory;
KW      vulnery; gene therapy; vaccine; gene; ss.
XX      XX
OS      Homo sapiens.
XX      XX
PN      WO2003042661-A2.
XX      XX
PD      22-MAY-2003.
XX      XX
PF      13-NOV-2002; 2002WO-US036810.
XX      XX
PR      13-NOV-2001; 2001US-0350666P.
PR      21-NOV-2001; 2001US-0332464P.
PR      29-NOV-2001; 2001US-0334393P.
PR      03-DEC-2001; 2001US-0335394P.
PR      14-DEC-2001; 2001US-0340376P.
PR      08-JAN-2002; 2002US-0347211P.
PR      10-JAN-2002; 2002US-0347349P.
PR      08-FEB-2002; 2002US-0355250P.
PR      13-FEB-2002; 2002US-0356714P.
PR      20-FEB-2002; 2002US-0359077P.
PR      29-MAR-2002; 2002US-0368099P.
PR      04-APR-2002; 2002US-0370110P.
PR      12-APR-2002; 2002US-0372246P.
PR      15-JUN-2002; 2002US-0386614P.
PR      16-JUN-2002; 2002US-0386839P.
PR      05-JUN-2002; 2002US-0393775P.
PR      22-JUL-2002; 2002US-0397845P.
PR      09-SEP-2002; 2002US-0409450P.
XX      XX
PA      (EOSB-) EOS BIOTECHNOLOGY INC.
XX      XX
PI      Afar D, Aziz N, Ginsburg NM, Gish KC, Glynn R, Hevezi PA;
PI      Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
DR      WPI; 2003-468649/44.
XX      XX
DR      P-PSDB; ADN39878.
XX      XX
PT      Determining the presence or absence of a pathological cell in a patient,
PT      useful for diagnosing, prognosing or treating cancer, comprises detecting
PT      a nucleic acid in a biological sample.
XX      XX
PS      Claim 8; SEQ ID NO C33; 1385pp; English.
XX      XX
SS      The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
XX      XX      whose expression is upregulated or downregulated in specific cancers or
XX      XX      other diseases such as angiogenic or fibrotic disorders, and to methods
XX      XX      of determining the presence or absence of a pathological cell in a
XX      XX      patient by detecting a nucleic acid at least 80% identical to those of
XX      XX      the invention or by detecting a polypeptide of the invention. The
XX      XX      invention also relates to expression vectors and host cells comprising a
XX      XX      nucleic acid of the invention; antibodies which specifically bind a
XX      XX      polypeptide of the invention; use of such antibodies for drug targeting;
XX      XX      and methods of screening for modulators of activity or expression of the
XX      XX      polypeptides and nucleic acids. The nucleic acids, polypeptides,
XX      XX      antibodies and methods are useful for diagnosing, prognosing and treating
XX      XX      cancer and other conditions such as psoriasis, ischaemia, heart disease,
XX      XX      atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
XX      XX      neovascularization syndromes, scarring and uterine fibroids. They may
XX      XX      also be useful in wound healing and in contraception. The present
XX      XX      sequence represents a nucleic acid sequence of the invention.
XX      XX
SQ      Sequence 1314 BP; 297 A; 369 C; 373 G; 275 T; 0 U; 0 Other;

```


Db 970 CTCTGATCATTTGATGGGCTTTACAGAGAGATGAGGAGATGTCTGACATACTG 1029
QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaTyrGln 361
Db 1030 CTGACAGGCTGATCCAGGATCATTCAGACACACGAGGTCATGACAGATGCGTACAG 1089
QY 362 G1yGlnValThrGluysMetMetCysAlaGly1LeProGlu1yG1yValAspThrCys 381
Db 1090 GGGAAAGTCCACGAAAGATGATGTGACGAGCATCCCGAAGGGGTGGACACCTGC 1149
QY 382 GlnG1yAspSerG1yG1yProLeuMetTyrGlnSerAspGlnThrIleValG1y1e 401
Db 1150 CAGGCTGACAGTGGTGGCCCTGATGATCCATCTGACAGTGCATGTGGTGGCATC 1209
QY 402 ValSerThrG1yTyrG1yCysG1yG1yProSerThrProG1yValTyrThrValSer 421
Db 1210 GTTAGCTGGGCGATGCTGCTGGGGCCCGAGACCCCGAGATATACCAAGGTCTCA 1269
QY 422 AlaTyrLeuAsnThrIleTyrAsnValITrplyAlaGluLeu 435
Db 1270 GCCTATCTCACTGATCTACATATGTCTGAAAGGCTGAGCTG 1311

RESULT 4

ADB80524
ID ADB80524 standard; DNA; 1314 BP.

AC ADB80524;

DT 04-DEC-2003 (first entry)

DE Ovarian cancer-associated transcript #45.

KW cytosolic; gene therapy; vaccine; ovarian cancer; diagnosis;

KW post-operative chemotherapy; radiation therapy; tumour prognosis;

KW pre-cancerous lesion detection; ds; gene.

OS Homo sapiens.

FN Key Location/Qualifiers

FT CDS 1..1314

FT /*tag= a

PN WO2002102235-A2.

PF 18-JUN-2002; 2002WO-US019297.

PR 18-JUN-2001; 2001US-0299234P.

PR 27-AUG-2001; 2001US-0315287P.

PR 05-SEP-2001; 2001US-0317544P.

PR 13-NOV-2001; 2001US-0350666P.

PR 12-APR-2002; 2002US-0372246P.

PA (BOSB-) BOS BIOTECHNOLOGY INC.

PI Mack DH, Gish KC;

DR WPI; 2003-167431/16.

DR P-PSDB; ADB80525.

PT Detecting an ovarian cancer-associated transcript in a cell from a

PT patient; comprises contacting a biological sample from the patient with a

PT polynucleotide that hybridizes to an ovarian cancer gene.

PS Claim 10; Page 305, 332pp; English.

XX The invention relates to a method of detecting an ovarian cancer-

XX associated transcript in a cell from a patient, by contacting a

XX biological sample from the patient with a polynucleotide that selectively

XX hybridizes to a sequence at least 80% identical to any of 80

XX nucleic acid sequences given in the specification. The method is useful

XX in diagnosing ovarian cancer and in identifying and using agents and/or

CC targets that inhibit ovarian cancer. The nucleic acid molecule,
CC polypeptide and the antibody may also be used in detecting ovarian
CC cancers, monitoring and early detection of relapse following treatment,
CC monitoring response to therapy, selecting patients for post-operative
CC chemotherapy or radiation therapy, in selecting mode of therapy,
CC determining tumour prognosis, early detection of pre-cancerous lesions,
CC and as vaccines. This sequence corresponds to one of the nucleic acids
CC used for the detection method of the invention.

XX
SQ Sequence 1314 BP; 297 A; 369 C; 373 G; 275 T; 0 U; 0 Other;

Alignment Scores:

Pred. NO.: 3.3e-207 Length: 1314
Score: 2337.00 Matches: 434
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.79% Indels: 0
DB: 10 Gaps: 0

US-10-803-530-2 (1-435) x ADB80524 (1-1314)

QY 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValIleProLeuAlaGlyLeuProArg 21

Db 10 GATCTGACAGTGCATCACTTGTACAGCTCTCGATGCAAAACCCCTGGCCAAACCCCGT 69

QY 22 IleProMetGluThrPheArgLysValG1y1LeProIleIleIleAlaLeuLeuSerLeu 41

Db 70 ATCCCCATGAGAACCTTCAGAAAGGTGGGAGATCCCATCATCATACATGACATCTGAGCTG 129

QY 42 AlaSerIleIleIleValValIleuValIleuValIleuValIleuValIleuValIleuVal 61

Db 130 GCGATATACATCATTTGTGTGCTTCTCATAGGTGATTTGTGATTAATCTACTTCTTC 189

QY 62 CysGlyGlnProLeuAsnIlePheArgLysValIleuValIleuValIleuValIleuValIleu 81

Db 190 TGGGGGAGCTCTCCATCTTCCATCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 249

QY 82 ProLeuGlyGluAspGlyGluHisCysValIleSerPheProGluGlyProAlaValAla 101

Db 250 CCTTGGGGGAG 309

QY 102 ValArgLeuSerIleAspArgSerThrLeuGlnValIleuAspSerIleThrGlyAsnTrp 121

Db 310 GTCCGCTCTCCAAAGAGCCATCCACAGAGTGTGAGACTCGGCCACAGGAACTGG 369

QY 122 PheSerIleCysPheAspAsnThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141

Db 370 TTCCTGCTGTTTCGACAACTTCACAGAGGCTCTGCGAGACACGCTGTAGGAGATG 429

QY 142 G1yTyrSerSerIleProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161

Db 430 GGTACAGCAGCAAAACCTTTCAGAGCTGTGGAGATTGGGCCACAGCATCTGGAT 489

QY 162 ValValGluIleThrGluAsnSerGlnIleuArgMetArgAsnSerSerGlyProCys 181

Db 490 GTTGTGAATATCAGAAACAGCCAGAGCTTCCATGCGAACTCAAGTGGCCCTGT 549

QY 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyIleSerLeuIleThrPro 201

Db 550 CTCTCAGGCTCCCTGCTCTCCCTGCACTGTCTTGGCTGTGGAGAGAGCTAAACCCCC 609

QY 202 ArgValValG1yG1yGlnGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221

Db 610 CGGTGTGGTGGTGGGAGAGGCTCTGTGATTTCTTGGCCCTGGCAGTACATCAG 669

QY 222 TyrAspLysGlnHisValCysG1yG1ySerIleLeuAspProHisSTrpyAlleuThrAla 241

Db 670 TACGCAAAACAGCAGCTGTGGAGAGAGATCTTGACCCCACTGAGTCTTACAGCA 729

QY 242 AlaHisCysPheAsnArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261

Db 730 GCCCACTGCTTCAAGAAACATACGATGTGTTCAATGGAAGGTCGGGACGGGTACAG 789

Db 1460 TCACCTATCTCAACTGATCTACATGTCTGGAAGCTGAGCTG 1504

RESULT 3

ABX76354
ID ABX76354 standard; DNA; 1314 BP.

AC ABX76354;

DT 02-APR-2003 (first entry)

XX Lung cancer-associated polynucleotide #218.

DE Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;
KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

OS Unidentified.

XX WO200286443-A2.

XX 31-OCT-2002.

PF 18-APR-2002; 2002WO-US012476.

XX 18-APR-2001; 2001US-0284770P.

PR 10-MAY-2001; 2001US-0290492P.

PR 09-NOV-2001; 2001US-0339245P.

PR 13-NOV-2001; 2001US-0350666P.

PR 29-NOV-2001; 2001US-0334370P.

PR 12-APR-2002; 2002US-0372246P.

XX (EOSB-) EOS BIOTECHNOLOGY INC.

XX Aziz N, Murray R;

XX WPI: 2003-093161/08.

DR P-PSDB; AB056625.

PT Detecting a lung cancer-associated transcript in a cell from a patient

PT for treating lung cancer, by contacting a biological sample from the

PT patient with a polynucleotide that exhibits increased or decreased

PT expression in lung cancer.

XX Claim 22; Page 353; 453bp; English.

XX The invention relates to a method for detecting a lung cancer-associated
CC transcript in a cell from a patient, comprising contacting a biological
CC sample from the patient with a polynucleotide that selectively hybridizes
CC to a sequence that is at least 80 % identical to a gene that exhibits
CC increased or decreased expression in lung cancer samples. Lung cancer-
CC associated polynucleotides and polypeptides are used for identifying a
CC compound that modulates a lung cancer-associated polypeptide, for
CC inhibiting proliferation of a lung cancer-associated cell to treat lung
CC cancer in a patient and for treating a mammal having lung cancer by
CC administering a modulatory compound identified. The methods are useful
CC for treating lung cancer, such as small cell lung cancer, non-small cell
CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
CC for diagnostic purposes and as targets for screening for therapeutic
CC compounds that modulate lung cancer, such as antibodies. Sequences
CC ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the
CC invention

XX SQ Sequence 1314 BP; 297 A; 369 C; 373 G; 275 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 3,3e-207 Length: 1314
Score: 2337.00 Matches: 434

Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 99.79%
DB: 8 Gaps: 0

US-10-803-530-2 (1-435) x ABX76354 (1-1314)

QY	2	AspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysProArg	21
DB	10	GATCTGACAGTATCAACTCTGAACAGCTCGATGATCAAAACCCCTCGCAAAACCCCGT	69
QY	22	IlleProMetGlnThrPheArgLysValGlyIleProIleIleIleValLeuLeuSerLeu	41
DB	70	ATCCCAATGAGACCTTCAGAAAGTGGGGATCCCATCATCATAGACACTAGAGCTTG	129
QY	42	AlaSerIleIleIleValValLeuIleLysValIleLeuAspLysIleValPheLeu	61
DB	130	GCGAGTATCATCATTTGGTGTGCTCATCAAGGTGATTCGATTAATACTACTTCTC	189
QY	62	CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGlnLeuAspCys	81
DB	190	TGCGGCGAGCTCTCCACTTTCATCCAGAGAGAGCTGTGTGACGAGAGCTGACTGT	249
QY	82	ProLeuGlyLysAspGlnLysCysValLysSerPheProGlyProAlaValAla	101
DB	250	CCCTTGGGGAGAGCAGAGACATGTGTCAAGACTTCCCGAAGGGCTGAGTGCA	309
QY	102	ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp	121
DB	310	GTCGCGCTCTCCAAAGGACCGATCCACACTGCAAGGTGCTGGAGCCGACACAGGAATCG	369
QY	122	PheSerAlaCysPheAspAsnPheThrGlnAlaLeuAlaGlnThrAlaCysArgGlnMet	141
DB	370	TTCCTGCTGTTCGACAACTTCACAGAACTTCGCTGAGACGCTGTAGGCAATG	429
QY	142	GlyTyrSerSerLysProThrPheArgAlaValGlnIleGlyProAspGlnAspLeuAsp	161
DB	430	GGTACAGACAGCAAAACCACTTCAGAGCTGTGAGATTGGCCACAGACTGTGAT	489
QY	162	ValValGlnIleThrGlnAsnSerGlnLeuArgMetArgAsnSerGlyProCys	181
DB	490	GTTCTTGAATACACAGAAACAGCAGAGCTTTCGATGCGCACTCAAGTGGCCCTGT	549
QY	182	LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuThrPro	201
DB	550	CTTCAGGCTCTCGTCTCTGCTGCACTGCTTTCCTGGAAGAGCTGAAGACCCCC	609
QY	202	ArgValValGlyGlyGlnLysLysValAspSerTrpProTrpGlnValSerIleGln	221
DB	610	CGTGTGAGGAGTGGGAGAGAGGCTCTGTGATTTCTTGAGCTTGAGGTGAGCATCAG	669
QY	222	TyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAla	241
DB	670	TACACAAACAGCAGCTGTGGAGGAGCATCTCGAACCCCACTGGGTCTTCACGGCA	729
QY	242	AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp	261
DB	730	GCCACGTCTTCAGAAACATACCGATGTGTTCACTGGAAGGTGGGAGGCTCAGAC	789
QY	262	LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleIleGlnPheAsnPro	281
DB	790	AAATGGGAGACCTCCATCTGCTGTGGCCCAAGATCATCATCATTTGAATTCACCCC	849
QY	282	MetTrpProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer	301
DB	850	ATGTACCCCAAAAGACATGACATGCGCTCATGTGAAGTGCAGTTCACACTTCCTTCA	909
QY	302	GlyThrValArgProIleCysLeuProPhePheAspGlnLysLeuThrProAlaThrPro	321
DB	910	GGCACAGTACAGGCCCATCTGTCTGCTTTTATATAGAGACTCATCCAGCAACCCCA	969
QY	322	LeuTrpIleIleGlyTyrPglyPheThrLysGlnAsnGlyLysMetSerAspIleLeu	341

XX Serine protease; human; HUPM-6; cell proliferation; cancer;
 KW immune disorder; inflammation; therapy; ss.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FT CDS 200..1507
 FT /tag= a
 XX MO936550-A2.
 XX 22-JUL-1999.
 XX 12-JAN-1999; 99WO-US000655.
 XX 16-JAN-1998; 98US-00008271.
 XX (INCY-) INCYTE PHARM INC.
 XX Bandman O, Hillman JL, Yue H, Guegler KU, Corley NC, Tang YT;
 PI Shah P;
 XX WPI, 1999-430616/36.
 DR P-PSDB; AAY06437.
 XX Novel human protease molecules useful in the treatment of developmental
 PT disorders and/or cancers.
 XX Claim 8, Page 86-87; 90pp; English.
 XX This nucleotide sequence codes for HUPM-6 (see AAY06437), a novel human
 CC protease. HUPM-6 cDNA was initially identified in Incyte Clone 1337018
 CC from the colon cDNA library COLNORT3 using a computer search for amino
 CC acid sequence alignments. The present sequence is a consensus sequence
 CC derived from overlapping and/or extended nucleic acid sequences: Incyte
 CC Clones 1271725 (TESTTUT02), 1337018, 586982 and 588598 (UTMSNOT01). A
 CC fragment comprising nucleotides 900-949 of the present sequence can be
 CC used for hybridisation. This sequence encompasses an active site residue.
 CC Northern analysis shows expression of HUPM-6 in gastrointestinal, and male
 CC an female reproductive cDNA libraries. Approximately 65% of these
 CC libraries are associated with neoplastic disorders, and 22% with the
 CC immune response. The invention provides 12 new human proteases, i.e. HUPM
 CC -1 to -12 (see AAY06437-43), and the polynucleotides encoding them (see
 CC AAX87149-60). Also provided are vectors, host cells and methods for
 CC producing HUPM polypeptides, as well as agonists and antagonists of HUPM.
 CC Methods for treating or preventing cell proliferative disorders and
 CC immune disorders using HUPM or HUPM antagonists are claimed
 XX
 SQ Sequence 2038 BP; 462 A; 591 C; 569 G; 416 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 4.76e-207 Length: 2038
 Score: 2338.00 Matches: 434
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.77% Mismatches: 0
 Query Match: 99.83% Indels: 0
 DB: 2 Gaps: 0
 US-10-803-530-2 (1-435) x AAX87154 (1-2038)
 QY 1 MetAspProAspSerAspGlnProLeuAAsenSerLeuAspValLeProLeuAArgLysPro 20
 Db 200 ATGATCTCTGACGATCATCACTCTGAAACGCTCGATGTAAACCCCTGGCMAACCC 259
 QY 21 ArgGleProMerGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSer 40
 Db 260 CGATCCCATGAGACCTTCAGAAAGGTGGGATCCCATCATCATGACACTACTAGC 319
 QY 41 LeuAlaSerIleIleIleValValValLeuIleLysValIleLeuAspLysTyrTyrPhe 60
 Db 320 CTGGCGAGATCATCTTGTGTGTCTCATCAAGGTGATTCGGATTAATACTACTTTC 379

QY 61 LeuCyGlyGlnProLeuHiIlePheProArgLysGlnLeuCyAspGlyGluLeuAsp 80
 Db 380 CTCTGGGGGAGAGCTCTCCACTTCATCCCGAAGAGACAGCTGTGACGAGAGCTGGAC 439
 QY 81 CysProLeuGlyGluAspGluGluHisCysValLysSerPheProGluGlyProAlaVal 100
 Db 440 TGTCCCTTGGGGAGAGACAGAGACAGCTGTGTCAAGACCTTCCGAAAGGGCTGGCACTG 499
 QY 101 AlaValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsn 120
 Db 500 GCAGTCCGCTCTCCAAAGACGATCCACACTGACAGTGTGGACTCGGCCACAGGAAAC 559
 QY 121 TrpPheSerAlaCysPheAspAsnThrThrGluAlaLeuAlaGluThrAlaCysArgGln 140
 Db 560 TGGTCTCTGCTGTGTTTGACAACTTCACAAAGCTTCGCTGAGACAGCTGTAGGAG 619
 QY 141 MetGlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGluAspLeu 160
 Db 620 ATGGGCTACAGCAGAAACCCACTTCAGACTGTGGAGATTGGCCCGACAGCAGATCTG 679
 QY 161 AspValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyPro 180
 Db 680 GATGTGTGTAATACAGAAACAGCCAGGAGCTTCGATGCGGAATCAAGTGGGCC 739
 QY 181 CysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThr 200
 Db 740 TGTCTCTAGGCTCTCTGTGTCTCTGCTGACTGTCTTCTGTGGGAGACCTGAAGACC 799
 QY 201 ProArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIle 220
 Db 800 CCCCGTGTGGTGGGGAGAGAGCTCTGTGTGATCTTGCTGGCTTGACAGCTCAGCATC 859
 QY 221 GlnTrpAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThr 240
 Db 860 CAGTACGACAAACAGCAGCTGTGGAGGAGATCTTGACCCCACTGGCTGTCTCAGC 919
 QY 241 AlaAlaHisCysPheAspGlyHisThrAspValPheAsnTrpLysValArgAlaGlySer 260
 Db 920 GCAGCCCACTGCTTCAGAGAAACATACGATGTGTTCACATCGAAGGTGGCGGAGGCTCA 979
 QY 261 AspLysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleIleGluPheAsn 280
 Db 980 GACAACTGGGACGCTCCCATCTCGGTGTGGCCAAAGTATCATCATGATTCAAC 1039
 QY 281 ProMetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPhe 300
 Db 1040 CCCATGTACCCCAAAAGCAATGACATCGCCCTCATGAAGCTGACATCCACTCATCTTC 1099
 QY 301 SerGlyThrValArgProIleCysLeuProPheAspGluGluLeuThrProAlaThr 320
 Db 1100 TCAGGCACACTCAGGCGCCATCTGTCTGCTTTGATGAGAGCTCATCTCACCCAC 1159
 QY 321 ProLeuTrpIleIleGlyTyrGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIle 340
 Db 1160 CCACTCGATGATCTTGTGATGGGCTTTACGAAGAAATGAGGGGAAGATGTCTACATA 1219
 QY 341 LeuLeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyr 360
 Db 1220 CTGCTGACAGCGCTCATCTCAGGTCAATGACACACACGCTGCAATGACAGACATCTGTAC 1279
 QY 361 GlnGlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThr 380
 Db 1280 CAGGGGGAAGTACCGAGAAATGATGTGTGACGACATCCCGAAGGGGGGTGTGACACC 1339
 QY 381 CysGlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGly 400
 Db 1340 TGCAGAGTGAAGAGTGTGGGCCCTGTGATGATCAATGTGACAGTGCATGTGTGGTGGC 1399
 QY 401 IleValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysVal 420
 Db 1400 ATGTTAGTGTGGGCTATGTGTGTGGGGGCCCGACACCCAGGAGTATACACCAAGTTC 1459
 QY 421 SerAlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435

4 102e

XX 04-JUL-2000; 2000WO-EP006211.
PF 12-JUL-1999; 99EP-00113448.
XX (MERE) MERCK PATENT GMBH.
XX Suendermann B, Hofmann U, Matzku S, Wilbert O;
PI WPI: 2001-147177/15.
DR P-PSDB; AAY72558.
XX New extracellular serine protease Seripancrin, useful for treating
PT cancer, osteoporosis, arthritis, chronic obstructive pulmonary disease,
PT diabetes, inflammatory disorders, stroke, angiogenesis and aberrant wound
PT healing.
XX Claim 5; Page 37-39; 45BP; English.
XX The present invention relates to seripancrin polynucleotides, and
CC polypeptides encoded by them. Seripancrin are members of serine protease
CC family. This protein contains a transmembrane domain, a low density
CC lipoprotein (LDL) domain, protease domain and a scavenger receptor
CC cysteine-rich (SRCR) domain. The LDL and SRCR domains help to define the
CC specificity of seripancrin's intra and intermolecular interactions. The
CC polynucleotides and polypeptides of the invention are useful for treating
CC and diagnosing diseases such as arthritis, chronic obstructive pulmonary
CC disorder (COPD), cancer, osteoporosis, aberrant wound healing,
CC angiogenesis, inflammatory disorders, diabetes, stroke and cardiovascular
CC diseases. Seripancrin genes are useful in chromosome localization
CC studies, as tools for tissue expression studies and also in gene therapy.
CC The polypeptides of the invention are used for identifying agonists and
CC antagonists useful for treating conditions associated with seripancrin
CC imbalance. These polypeptides are also useful as vaccines. The present
CC sequence is a cDNA coding for seripancrin protein. The seripancrin gene
CC is located on human chromosome 11q22-q23
XX
SQ Sequence 1305 BP; 293 A; 368 C; 372 G; 272 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1.12e-207 Length: 1305
Score: 2342.00 Matches: 435
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-10-803-530-2 (1-435) x AAD02556 (1-1305)
QY 1 MetAspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysPro 20
DB 1 ATGGATCTGACAGATCAACCTCTGAACAGCTCGATGTCMAACCCCTGGGCAAAACC 60
QY 21 ArgIleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSer 40
DB 61 CGTATCCCAAGGAGACCTTCAGAAAGGTGGGATCCCATCATATAGCACTACTAGGC 120
QY 41 LeuAlaSerIleIleIleValValValLeuIleLysValIleLeuAspLysTyrTrpHe 60
DB 121 CTGGGAGTATCATCATTTGTGTCTCTCATCAAGTGTGATTCGATTAATACTACTTC 180
QY 61 LeuGlyGlyGlnProLeuHisPheIleProArgLysGlnLeuGlyAspGlyGlnLeuAsp 80
DB 181 CTCTGGGGGAGCCCTCTCACTTCATCCGAGGAAGCAGCTGTGTGACGGAGAGCTGAGC 240
QY 81 CysProLeuGlyGlnLysAspGlyGlnLysIleCysValLysSerPheProGlyGlyProAlaVal 100
DB 241 TGTCTCTTGGGGGAG 300
QY 101 AlaValAlaGluSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsn 120
DB 301 GCAGTCCGCTCTCCMAAGAGCCGATCCACATGAGGCTGAGCTCGGCCACAGGGAAC 360

QY 121 TrpPheSerAlaCysPheAspAsnPheThrGlnAlaLeuAlaGluThrAlaCysArgGln 140
DB 361 TGTCTCTCTGCTCTTTTGCACAACTTCACAGAACTCTCGCTGAGAGAGCTGTAGGACAG 420
QY 141 MetGlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeu 160
DB 421 ATGGGCTTACAGACACAAACCCATTTCAAGCTGTGAGATGGCCACAGACAGATCTG 480
QY 161 AspValAlaGluIleThrGlnAsnSerGlnGluLeuArgMetArgAsnSerSerGlyPro 180
DB 481 GATGTGTGAATACAGAAACAGACGAGAGCTTGATGATGCGAACTCAAGTGGGCC 540
QY 181 CysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThr 200
DB 541 TGTCTCTCAGGCTCCCTGGTCTCCCTGCATGCTTTCCTTGAGGAAGGCTTGAAGCC 600
QY 201 ProArgValValGlyGlyGlnLysValAspSerTrpProTrpGlnValSerIle 220
DB 601 CCCCCTGTGGTGGGGAGAGAGCCCTGTGTGATTTCTGGCCCTGGCAGGTACGATC 660
QY 221 GlnTyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThr 240
DB 661 CAGTACGACAAACAGACAGTCTGTGAGGGAGCATCTGAGCCCTGGAGCTCTTCAAG 720
QY 241 AlaAlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySer 260
DB 721 GCAGCCCACTGCTTCAGGAAACATACCGATGTCTTCAACTGGAAAGTCCGGGAGGCTCA 780
QY 261 AspLysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsn 280
DB 781 GACAACTGGGACAGCTTCCATCCCTGCTGGCCCAAGATCATCATCATGAAATTCAAC 840
QY 281 ProMetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrHe 300
DB 841 CCCATGTAACCCCAAGCAAGATGAGCTCTCATGAAAGCTGACGTTCCCATCTTC 900
QY 301 SerGlyThrValAspProIleCysLeuProPhePheAspGlnGluLeuThrProAlaThr 320
DB 901 TCAGGACAGTCAAGGCCATGTGTCTGCTCTTTGATGAGAGCTCATCTCCAGCCACC 960
QY 321 ProLeuTrpIleIleGlyTyrGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIle 340
DB 961 CCACCTGTGATCATTTGATGGGGCTTTAAGAAAGAGATGAGATGTCGACATA 1020
QY 341 LeuLeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaThr 360
DB 1021 CTGCTGACAGGCGTCAAGTCCAGTCAATGACAGCACAGGTGCATATGACAGATGCCATC 1080
QY 361 GlnGlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThr 380
DB 1081 CAGGGGAGATCACCGAGAGATATGTGTCCAGGCATCCCGAAAGGGGTGTGGACACC 1140
QY 381 CysGlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGly 400
DB 1141 TGCCAGGGGTGACAGTGGGGGCCCTGTATGACCAATGTACAGTGCATGTGGTGGGC 1200
QY 401 IleValSerTrpGlyTyrGlyCysGlyGlyProSerThrTrpProGlyValTyrThrLysVal 420
DB 1201 ATCGTTAGCTGGGCTATGTGTGGGGGCCCGAGCACCCAGAGATATACCAAGAGTTC 1260
QY 421 SerAlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
DB 1261 TCAAGCTATCTCAACTGGATTTACAAATGTCTGAAAGGCTGAGCTG 1305
RESULT 2
ID AAX87154 standard; cDNA; 2038 BP.
XX AAX87154;
AC AAX87154;
XX 27-SEP-1999 (First entry)
DT
XX Human protease HUPM-6 cDNA.
DE

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: December 6, 2005, 18:45:35 / Search time 680 Seconds

(without alignments)
4263.443 Million cell updates/sec

Title: US-10-803-530-2

Perfect score: 2342
Sequence: 1 MPPDDQPLNSLDVKKPLRKP.....VYTKVSAYLNIYVWYKAEI 435

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus.p2n.model -DEV=xip
-O=/cgn2.1/USPTO_spool_P/US10803530/runat_05122005_083215_2111/app_query.fasta_1.583
-DB=N_Geneseq -QFMT=fastag -SUFFIX=ring -MINMATCH=0 -1 -LOOEXT=0 -LOOEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blom62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NORM=ext -HAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10803530 @CGN_1_1_1096 @runat_05122005_083215_2111 -NCPU=6 -ICPU=3
-NO_MMAP -LARGESUBSTRY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_21: *
1: geneseqn1980s: *
2: geneseqn1990s: *
3: geneseqn2000s: *
4: geneseqn2001as: *
5: geneseqn2001bs: *
6: geneseqn2002as: *
7: geneseqn2002bs: *
8: geneseqn2003as: *
9: geneseqn2003bs: *
10: geneseqn2003cs: *
11: geneseqn2003ds: *
12: geneseqn2004as: *
13: geneseqn2004bs: *
14: geneseqn2005s: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2342	100.0	1305	4	AAD02556 Human ser
2	2338	99.8	2038	2	AAX87154 Human pro
3	2337	99.8	1314	8	ABX76354 Lung can
4	2337	99.8	1314	10	ADB80524 Ovarian c

5	2337	99.8	1314	11	ADN39512 Cancer/an
6	2337	99.8	1314	11	ADN39661 Cancer/an
7	2337	99.8	1314	11	ADN39472 Cancer/an
8	2337	99.8	1314	11	ADN38815 Cancer/an
9	2337	99.8	1314	11	ADN39460 Cancer/an
10	2337	99.8	1314	13	ADU25739 CDNA enco
11	2337	99.8	2081	6	AA172976 CVA8 pref
12	2337	99.8	2104	12	ADM04863 Antipsori
13	2337	99.8	2104	14	ADM04403 Human Ova
14	2337	99.8	2104	14	AE54844 Human TAT
15	2337	99.8	2104	14	AE54844 Human TAT
16	2337	99.8	2104	14	AE54844 Human TAT
17	2335	99.7	2307	8	ACF12937 Human cer
18	2335	99.6	2590	13	ABZ35523 Human cer
19	2329	99.4	1479	6	AA164284 Human cer
20	2329	99.4	1479	6	ABZ35523 Human cer
21	2328	99.4	2627	13	ABD32788 Human cer
22	2324	99.2	2070	3	AA290471 Human cer
23	2324	99.2	2079	4	AA290471 Human cer
24	2324	99.2	2079	4	AA290471 Human cer
25	2324	99.2	2079	6	AA172975 Human cer
26	2324	99.2	2079	9	ACF36041 Human cer
27	2324	99.2	2079	10	ADE48026 DNA enco
28	2324	99.2	2079	10	ABSS7763 Human cer
29	2324	99.2	2079	12	ADG08800 Transemb
30	2324	99.2	2079	12	ADG08800 Transemb
31	2324	99.2	2079	12	ADG08800 Transemb
32	2324	99.2	2079	13	ADG08800 Transemb
33	2319	99.0	2137	14	AEBS5225 Human cer
34	2319	99.0	2137	10	AD110372 Human cer
35	2319	99.0	2137	12	ADJ46896 Human cer
36	2297.5	98.1	2063	3	AA337099 Human PRO
37	2297.5	98.1	2063	4	AA543396 DNA enco
38	2297.5	98.1	2063	4	AA543396 DNA enco
39	2297.5	98.1	2063	4	AA543396 DNA enco
40	2297.5	98.1	2063	6	AA543396 DNA enco
41	2297.5	98.1	2063	6	ABST7433 Human PRO
42	2297.5	98.1	2063	8	ABK11090 CDNA enco
43	2297.5	98.1	2063	8	ACA89539 CDNA enco
44	2297.5	98.1	2063	8	ACA73549 Human sec
45	2297.5	98.1	2063	8	ACA05864 Human sec

ALIGNMENTS

RESULT 1	
AAD02556	standard; cDNA, 1305 BP.
XX	
XX	
AC	AAD02556;
XX	
DT	02-MAY-2001 (first entry)
XX	
DE	Human seripancrin cDNA.
XX	
KW	Human; seripancrin; serine protease; chromosome 11q22-q23; therapy;
KW	arthritis; chronic obstructive pulmonary disorder; COPD; cancer;
KW	osteoporosis; aberrant wound healing; angiogenesis; diabetes;
KW	inflammatory disorder; stroke; cardiovascular disease; gene therapy;
XX	vaccine; cytosolic; cerebroprotective; vulnerary; osteopathic; ss.
OS	Homo sapiens.
XX	
PH	
FT	Key
FT	CDS
FT	Location/Qualifiers
FT	1..1305
FT	/*tag= a
FT	/product= "Human seripancrin protein"
FT	/note= "The coding region does not include stop codon"
FT	/partial
PN	WO200104141-A2.
XX	
PD	18-JAN-2001.

```

Db      1304 CTGCAGCGCTGATCCAGATCATTCAGACACACGGTGCATTGCAGATGCCGTACCAG 1363
Qy      362  |||||ThrgluValMetMetCysAlaGlyIleProGluGlyValAspThrCys 381
Db      1364 GGGGAAGTCACCGAAGATGATGTGCAGGCATCCGGAAGGGGGTGTGGACACCTGC 1423
Qy      382  GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle 401
Db      1424 CAGGCTGACAGTGGTGGGCCCTGATGTACCAATCTGACAGTGGCATGTGTGGGCATC 1483
Qy      402  ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSer 421
Db      1484 GTTAGCTGGGGCTATGCTGGGGGGCCGAGACCCCGAGAGTATACACCAAGGCTCTCA 1543
Qy      422  AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db      1544 GCCTATCTCAACTGATCTACAAATGTCTGGAAGGCTGAGCTG 1585

```

Search completed: December 7, 2005, 01:13:14
 Job time : 932 secs

Db 1202 CTCGGAATCATGATGGGGCTTTACGAGCAGATGAGGAGGATGCTGACATACCTG 1261
Qy 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGln 361
Db 1262 CTGAGGGGTGATGCTCCAGGTATTGACGACACGGTGCAATGACAGATGCGTACCAG 1321
Qy 362 GlyIuValThrGluYsMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys 381
Db 1322 GGGGAGACGACCGAAGATATGTGTCCAGGCATCCCGAAGGGGGTGTGACACCTGC 1381
Qy 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTyrPheIleValGlyIle 401
Db 1382 CAGGTGACAGTGGGGGCTCCCTGATGTAACCATCTGACAGTGCATGTGGGCTAC 1441
Qy 402 ValSerTrpGlyTyrGlyGlyCysGlyIleProSerThrProGlyValAlaTyrThrValSer 421
Db 1442 GTTAGCTGGGGCTATGCTGGGGGGCCGAGCACCCGAGATATACCAAGGCTTCA 1501
Qy 422 AlaTyrLeuAsnTrpIleTyrAsnValTyrPheAlaGluLeu 435
Db 1502 GCCTATCTCAACTGATCTACAAATGTCTGAAAGGCTGAGCTG 1543
RESULT 14
US-10-097-340-317
; Sequence 317, Application US/10097340
; Publication No. US20030087250A1
; GENERAL INFORMATION:
; APPLICANT: John MONAHAN
; APPLICANT: Manjula GANNAVARAPU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangi KAMATKAR
; APPLICANT: Steve G KOVARS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISSEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Peter VEIRY
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. BAST, JR.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHMANDT
; APPLICANT: Xumei ZHAO
; APPLICANT: Karen GLATT
TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer
FILE REFERENCE: MRI-030
CURRENT APPLICATION NUMBER: US/10/097,340
PRIOR FILING DATE: 2002-03-14
PRIOR APPLICATION NUMBER: 60/276,025
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/325,149
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/276,026
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/324,967
PRIOR FILING DATE: 2001/09/26
PRIOR APPLICATION NUMBER: 60/311,732
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: 60/325,102
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/323,580
PRIOR FILING DATE: 2001-09-19
NUMBER OF SEQ ID NOS: 363
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 317
LENGTH: 2307
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(2307)
OTHER INFORMATION: n = A,T,C or G

US-10-097-340-317
Alignment Scores:
Pred. No.: 1,8e-282 Length: 2307
Score: 2337.00 Matches: 434
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.79% Indels: 0
DB: 5 Gaps: 0
US-10-803-530-2 (1-435) x US-10-097-340-317 (1-2307)
Qy 2 AspProAspSerAspGlnProLeuAsnSerIleuAspValIleProLeuAlaGlyProArg 21
Db 284 GATCCTGACAGTATCAACCTCTGAACAGCCTCGATGTCAAAACCCCTGCGCAACCCCGT 343
Qy 22 IleProMetGluThrPheAlaGlyValGlyIleProIleIleIleAlaLeuLeuSerLeu 41
Db 344 ATCCCATGGAGACCTTCAGAAAGTGGGATCCCATCATCATATAGACTAGAGCTCG 403
Qy 42 AlaSerIleIleIleValIleValIleuIleIleValIleLeuAspLysTyrTyrPheLeu 61
Db 404 GCGAGTATCATATTGTTGTTTCTCTCATCAAGTGATTCGGATTAAATACTACTCTTC 463
Qy 62 CysGlyGlnProLeuHisPheIleProArgLysGlnIleuCysAspGlyGluLeuAspCys 81
Db 464 TGGGGGAGGCTCTCCACTTCATCCCGAGGAGAGCTGTGTGACGAGAGAGCTGACTGT 523
Qy 82 ProLeuGlyIuAspGluGluHisCysValIleSerPheProGluGlyProAlaValAla 101
Db 524 CCGTGGGGGAGGACGAGAGCAGTGTGTCAAGGCTTCCCGAAGGGCCCTGACAGTGC 583
Qy 102 ValArgLeuSerLysAspArgSerThrLeuGlnValIleuAspSerAlaThrGlyAsnTyr 121
Db 584 GTCCGCTCTCAAGACCAATCCATCCAGAGGTGTGACTGGCCACAGGAACTCG 643
Qy 122 PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
Db 644 TTCCTGTGCTGTTTGCACAACTTCACAGAACTCTCGGTGAGACGCTGTAGGACAGATG 703
Qy 142 GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
Db 704 GGTACAGACAGCAAAACCACTTTCAGAGCTGTGAGATGTGCCACAGACCATGTGAT 763
Qy 162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
Db 764 GTTGTGTAATCACAGAAACAGCCAGAGCTTGTGCAGCGGAACTCAAGTGGGCTGT 823
Qy 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro 201
Db 824 CTCACAGCTCCCTGGTCTCCCTGCACACTGTCTTGGCTGTGGAAAGGCTTAAGACCC 883
Qy 202 ArgValValGlyGlyGluGluAlaSerValAspSerTrpProTyrGlnValSerIleGln 221
Db 884 CGTGGGTGGGTGGGAGAGAGGCTCTGTGATCTTGGCTTGGCAGTGCAGCTCAG 943
Qy 222 TyrAspLysGlnHisValCysGlyIleSerIleLeuAspProHisIleThrValLeuThrAla 241
Db 944 TACGCAAAACAGACGCTGTGGAGGAGCATCTCGAGCCCACTGGGTCTTCACGGCA 1003
Qy 242 AlaHisCysPheAspGlySerHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
Db 1004 GCCACCTCTTCAGGAACATACCGATGTGTTCACCTGAAGGTGCGGGAGGCTCAAC 1063
Qy 262 LysLeuGlySerPheProSerLeuAlaValAlaIleIleIleIleGluPheAsnPro 281
Db 1064 AAACGGGAGCTTCCATCCCTGGCTGTGGCAAGATCATCATCATTAATTCAACCCC 1123
Qy 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db 1124 ATGTACCCCAAGAAAGATGACATGCGCTCATGAAAGCTGACAGTTCCTCACTTTC 1183
Qy 302 GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro 321

QY 222 TyrAspIysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValIleuThrAla 241
Db 902 TACACAAAGAGCGATGTGTGGAGGAGCATCTCGAACCCCGATGGGTCTTCAGCGCA 961
QY 242 AlaHisCysPheArgLysHisTrpAspValPheLeuTrpIysValAlaGlyIleSerAsp 261
Db 962 GCCACCTGCTTCAGGAAACATACGATGTGTTCACTGAAAGGGGCGGAGGCTCGAC 1021
QY 262 LysLeuGlySerPheProSerIleuAlaValAlaLysIleIleIleGluPheAspPro 281
Db 1022 AAACCTGGGAGCTTCCCATCCCTGGCTGTGGCCAAAGATCATCATTTGAATTCACCCC 1081
QY 282 MetTyrProIysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db 1082 ATGTACCCCAAGACAATGATGATGCCCTCATGTGAAGCTGACATTCACACTTCTCA 1141
QY 302 GlyThrValArgProIleCysIleuProPheAspGluGluLeuThrProAlaThrPro 321
Db 1142 GGCACAGTCAGGCCCATCTGTGCCCCCTTCTTGATGAGAGCTCACTCCAGCCACCCCA 1201
QY 322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyLysMetSerAspIleLeu 341
Db 1202 CTCGTGATCATTTGGATGGGCTTTAAGAACAGATGAGAGAAATGTCTGACATCTG 1261
QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGln 361
Db 1262 CTCACAGCGTCAGTCCAGGTCAATTCAGACACACGCGTGCATGACAGATCGCTACAG 1321
QY 362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyValAspThrCys 381
Db 1322 GGGGAAGTCACCGAAGATGATGTGTGCGAGCATCCCGAAGGGGTGTGGACACCTGC 1381
QY 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValIleGlyIle 401
Db 1382 CAGGGTGACAGTGTGGGCCCCGTGATGTACCAATTCAGACAGTGGCATGTGTGGGCAATC 1441
QY 402 ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValIleThrLysValSer 421
Db 1442 GTTAGCTGGGGCTATAGCTGTGGGGGCGCCGACGACCCGAGAGTATATACCAAGGTCTCA 1501
QY 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db 1502 GCGTATCTCACTGATCTTCAATGTCTGGAGGCTGAGCTG 1543
RESULT 13
US-10-991-287-2
/ Sequence 2, Application US/10991287
/ Publication No. US20050208523A1
/ GENERAL INFORMATION:
/ APPLICANT: Ashkenazi, Avi J.
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Guiney, Austin L.
/ APPLICANT: Polakis, Paul
/ APPLICANT: Smith, Victoria
/ APPLICANT: Wood, William I.
/ APPLICANT: Wu, Thomas D.
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
/ FILE OF INVENTION: TREATMENT OF TUMOR
/ FILE REFERENCE: P5037R1-US
/ CURRENT APPLICATION NUMBER: US/10/991,287
/ PRIOR APPLICATION NUMBER: 2004-11-17
/ PRIOR FILING DATE: 2003-11-20
/ NUMBER OF SEQ ID NOS: 10
/ SEQ ID NO 2
/ LENGTH: 2104
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-991-287-2
Alignment Scores: 1.56e-282 Length: 2104
Pred. No.: 1

Score: 2337.00 Matches: 434
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.79% Indels: 0
DB: 9 Gaps: 0
US-10-803-530-2 (1-435) x US-10-991-287-2 (1-2104)
QY 2 AspProAspSerAspGlnProLeuAsnSerIleuAspValIysProLeuArgLysProArg 21
Db 242 GATCTGACAGATGATCACTTCTGAAACAGCTTCGATGTCMAACCCCTGGCAAAACCCGT 301
QY 22 IleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSerLeu 41
Db 302 ATCCCATGGAACCTTCAGAAAGGTGGGAGATCCCATCATATATAGCATCTGAGCTTG 361
QY 42 AlaSerIleIleIleValIleValIleuIleLysValIleLeuAspLysTyrThrPheLeu 61
Db 362 GCGAGATCATCATTTGGTGTCTCTCATCAAGGTGATTCGTGATTAATTAATACTTCTGC 421
QY 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys 81
Db 422 TCGGGCAGCTCTCTCACTTCAATCCCGAAGACAGCTGTGTGACGAGAGCTGACCTGT 481
QY 82 ProLeuGlyGluAspGluGluHisCysValLysSerPheProGluGlyProAlaValAla 101
Db 482 CCTTGGGGGAGAGAGAGAGACATGTCTCAAGACTTCCCGAAGGCTCTGACGTGCA 541
QY 102 ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp 121
Db 542 GTCCGCTCTCCMAAGACCGATCCACATGACAGGTGTGACTCGGCACAGGGAATCTGG 601
QY 122 PheSerAlaCysPheAspAsnPheThrGlnAlaLeuAlaGluTrpAlaCysArgGlnMet 141
Db 602 TTCTTCTGCTGTTTGAACACTTTCACAGAGCTCTCGTGGAGACAGCTGTAGGACATG 661
QY 142 GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
Db 662 GGCCTACAGAGAAACCACTTTCAGAGCTGTGAGATTTGGCCAGACCAAGATCTGGAT 721
QY 162 ValValGluIleThrGlnAsnSerGlnLeuArgMetArgAsnSerSerGlyProCys 181
Db 722 GTTGTGAATCACAGAAACAGCCAGAGACTTGCAATGCGGAATCAAGTGGGCCCTGT 781
QY 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuThrPro 201
Db 782 CTCCTAGGCTCCCTGGTCTCCCTGACTGTCTGTGTGGAAAGGCTGAAAGCCCCC 841
QY 202 ArgValValIleGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
Db 842 CGTGGTGGGTGGGGAGAGAGGCTGTGTGATTTCTTGGCTTGGCAGGTACGATCCAG 901
QY 222 TyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValIleuThrAla 241
Db 902 TACGACAAAGCAGCGTGTGTGAGGAGCATCTCGAACCCCACTGGGTCTTCACGCGCA 961
QY 242 AlaHisCysPheArgLysHisTrpAspValPheLeuTrpIysValAlaGlyIleSerAsp 261
Db 962 GCCACCTGCTTCAGGAAACATACGATGTGTTCACTGAAAGGGGCGGAGGCTCGAC 1021
QY 262 LysLeuGlySerPheProSerIleuAlaValAlaLysIleIleIleGluPheAspPro 281
Db 1022 AAACCTGGGAGCTTCCCATCCCTGGCTGTGGCCAAAGATCATCATTTGAATTCACCCC 1081
QY 282 MetTyrProIysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db 1082 ATGTACCCCAAGACAATGATGATGCCCTCATGTGAAGCTGACATTCACACTTCTCA 1141
QY 302 GlyThrValArgProIleCysIleuProPheAspGluGluLeuThrProAlaThrPro 321
Db 1142 GGCACAGTCAGGCCCATCTGTGCCCCCTTCTTGATGAGAGCTCACTCCAGCCACCCCA 1201
QY 322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyLysMetSerAspIleLeu 341

```
OY 122 PheSerAlaCySPheASPsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
Db 602 TTCTCTGCTGCTTTTCGACACTTTCACAGAGCTCTCGTGAGACAGCTGTAGGCAAGT 661
OY 142 G1YrSerSerLysProThrPheArgAlaValGluIleGlyProASPGLnASPLeuASP 161
Db 662 GGCTACAGACAGAAACCCACTTTTCAGAGCTGTGGAGATGTGGCCAGACAGAGATCTGAT 721
OY 162 ValValGluIleThrGluASPsnSerGlnLeuLeuArgMetArgASPsnSerGlyProCys 181
Db 722 GTTGTGTAATACACAGAAACAGCCAGAGCTTCCCATCGCGAATCTCAAGTGGGCCCTGT 781
OY 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro 201
Db 782 CTCACAGGCTCCCTGGTCTCCCTGCATGCTCTTGGAGAAAGCTGAGAGCCCCC 841
OY 202 ArgValValGlyGlyGluGluAlaSerValASPserTPPProTPRGLValSerIleGln 221
Db 842 CGTGTGGTGGTGGGAGAGAGGCTCTGTGATTTTGCCCTTGGCAGGTACGATCCAG 901
OY 222 TyrASPlySGlnHisValCysGlyGlySerIleLeuASPProHisTPRValLeuThrAla 241
Db 902 TACGACAAACAGACAGTCTGTGGAGGAGCATCTGACCCCACTGGGTCCTCACGGCA 961
OY 242 AlaHisCysPheArgLysHisThrASPValPheASPnTPRlyValArgAlaGlySerASP 261
Db 962 GCCCACTCTTCAGAAACATACCATGTCGTTCACTGAAAGTCCGGGCGAGGCTCAGAC 1021
OY 262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheASPnPro 281
Db 1022 AAACGGGAGAGCTTCCATCCCTGGCTGTGGCAAGATCATCATATTGAATTCACCCC 1081
OY 282 MetTyrProLysASPsnASPnIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db 1082 ATGTACCCCAAGACATGACATGCCCTCATGAAGCTGCAGTTCCCATCTTCTTCA 1141
OY 302 GlyThrValArgProIleCysLeuBProPhePheASPGLnGluLeuThrProAlaThrPro 321
Db 1142 GGCAAGTCAAGGCCCATCTGTCTGCCCTTCTTGTAGAGAGCATCTCCAGCCACCCCA 1201
OY 322 LeuTPRleIleGlyTyrGlyPheThrLysGlnASPnGlyLysMetSerASPnIleLeu 341
Db 1202 CTCTGGATCATTTGATGGGCTTTTACGAAGCATGGAGGAGATGTCTGACACTACTG 1261
OY 342 LeuGlnAlaSerValGlnValIleASPserThrArgCysASPnAlaASPnAlaTyrGln 361
Db 1262 CTGCGGGGCTCAGTCAAGTCAATTGACAGACACAGGTGCAATGACAGATGCTTAC 1321
OY 362 GlyGluValThrGluLysMetCysAlaGlyIleProGluGlyValValASPnThrCys 381
Db 1322 GGGGAGTCAACGGAAGAATGATGTGTGAGGCACTCCGGAAGGGGTGTGACACCTGC 1381
OY 382 GlnGlyASPserGlyGlyProLeuMetTyrGlnSerASPnTPRHisValValIle 401
Db 1382 CAGGCTGACAGTGTGGGCCCTGTGATGCCAATCTGACCACTGAGTGTGGGATC 1441
OY 402 ValSerTPRGLyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSer 421
Db 1442 GTTACTGGGGCTAAGCTGCGGGGGCCGAGACACCCGAGGATATACCAAGGTCTCA 1501
OY 422 AlaTyrLeuASPnTPRleTyrASPnValTPRlySPnAlaGluLeu 435
Db 1502 GCCTATCTCACTGATCTACAAATGTCTGGAAAGGCTGAGCTG 1543
```

RESULT 12
US-10-994-117-2

; Sequence 2, Application US/10994117
; Publication No. US20050164250A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.

```
; APPLICANT: Polakis, Paul  
; APPLICANT: Smith, Victoria  
; APPLICANT: Wood, William I.  
; APPLICANT: Wu, Thomas D.  
; APPLICANT: Zhang, Dong-Xiao  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND  
; FILE REFERENCE: P503782  
; CURRENT APPLICATION NUMBER: US/10/994,117  
; PRIOR FILING DATE: 2004-11-19  
; PRIOR FILING DATE: 2004-11-20  
; NUMBER OF SEQ ID NOS: 10  
; SEQ ID NO 2  
; LENGTH: 2104  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-994-117-2  
Alignment Scores:  
Pred. No.: 1,56e-282 Length: 2104  
Score: 2337.00 Matches: 434  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 99.79% Indels: 0  
Gaps: 0  
DB: 9  
US-10-803-530-2 (1-435) x US-10-994-117-2 (1-2104)  
OY 2 ASPProASPserASPGLnProLeuASPsnSerLeuASPValLysProLeuArgLysProArg 21  
Db 242 GATCTGACATGATCAACCTTCTGACAGAGCTTCATGATCAAAACCCCTGGCAAAACCCCT 301  
OY 22 IleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSerLeu 41  
Db 302 ATCCCATGAGACCTTCAGAAAGTGGGATCCCATCATCATACATGACATCTAGCCTG 361  
OY 42 AlaSerIleIleIleValValLeuIleLysValIleLeuASPnLysTyrPheLeu 61  
Db 362 GCGAGTATCATATTGTGTGTCTCTCATCAAGGTGATTCGATTAATACTACTTCTC 421  
OY 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysASPnGlyLeuASPnCys 81  
Db 422 TGGCGGAGCTCTCCACTTCATCCAGAGAGAGCTGTGTGAGGAGAGAGCTGACTGT 481  
OY 82 ProLeuGlyLysASPnArgSerThrLeuGlnValLeuASPsnAlaThrGlyASPnTPR 101  
Db 482 CCTTTGGGGAGACAGAGACATCTGTCTCAAGACTTCCCGAAGGCTGTGAGTGGCA 541  
OY 102 ValArgLeuSerLysASPnArgSerThrLeuGlnValLeuASPsnAlaThrGlyASPnTPR 121  
Db 542 GTCCGCTCTCCAGAGACCATTCACACTGAGAGTGTGACTGGGCCACAGGGAATCG 601  
OY 122 PheSerAlaCysPheASPsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141  
Db 602 TTCTCTGCTGCTTTTCGACACTTTCACAGAGCTCTCGTGAGACAGCTGTAGGCAAGT 661  
OY 142 G1YrSerSerLysProThrPheArgAlaValGluIleGlyProASPGLnASPLeuASP 161  
Db 662 GGCTACAGACAGAAACCCACTTTTCAGAGCTGTGGAGATGTGGCCAGACAGAGATCTGAT 721  
OY 162 ValValGluIleThrGluASPsnSerGlnLeuLeuArgMetArgASPsnSerGlyProCys 181  
Db 722 GTTGTGTAATACACAGAAACAGCCAGAGCTTCCCATCGCGAATCTCAAGTGGGCCCTGT 781  
OY 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro 201  
Db 782 CTCACAGGCTCCCTGGTCTCCCTGCATGCTCTTGGAGAAAGCTGAGAGCCCCC 841  
OY 202 ArgValValGlyGlyGluGluAlaSerValASPserTPPProTPRGLValSerIleGln 221  
Db 842 CGTGTGGTGGTGGGAGAGAGGCTCTGTGATTTTGCCCTTGGCAGGTACGATCCAG 901
```

Query Match: 99.79% Indels: 0
 DB: 3 Gaps: 0
 US-10-803-530-2 (1-435) x US-09-851-588-7 (1-2081)

QY 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValIysProLeuArgIysProArg 21
 DB 224 GATCTGACAGTGAATCACTCTGAAACAGCTCCGATGCAAAACCCCTGCGAAACCCCGT 283
 QY 22 IlePrometGluThrPheArgIysValGlyIleProIleIleIleIleAlaLeuLeuSerLeu 41
 DB 284 ATCCCAATGAGAGACCTTCAAGAAAGGTGGGATCCCATCATCATAGACATCACTGAGCTG 343
 QY 42 AlaserIleIleIleValIleValIleValIleValIleValIleValIleValIleValIle 61
 DB 344 GCGATATCATCATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 403
 QY 62 CysGlyGlnProLeuHisPheIleProArgIysGlnLeuCysAspGlyGluLeuAspCys 81
 DB 404 TGGGAGAGCTCTCCACTTCATCCGAGAAAGCAGCTGTGTGACGAGAGCTGTGACTGT 463
 QY 82 ProLeuGlyGluAspGlnGluHisCysValIysSerPheProGluGluProAlaValAla 101
 DB 464 CCCTTGGGGAGAGACGAGAGACCTGTGTCAAGAGCTTCCCGAAAGGCTGTGACGTGGCA 523
 QY 102 ValArgLeuSerIysAspArgSerThrLeuGlnValIleuAspSerAlaThrGlyAsnTrp 121
 DB 524 GTCCGCTCTCCAGAGACCGATCCACCTGAGGTGTGTGACTGCGCCACAGAGAACTGG 583
 QY 122 PheSerIaCysPheAspAsnPheThrGluAlaLeuAlaGluThrIaCysArgGlnMet 141
 DB 584 TTCTCTGCTGTTTGCACCACTTCAAGAGCTCTGCTGAGACAGCTGTAGAGAGAG 643
 QY 142 GlyTyrSerSerIysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
 DB 644 GCGTACGAGCAAAACCACTTTCAGAGCTGTGAGATGGCCCAACAGAGATCTGGAT 703
 QY 162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
 DB 704 GTTGTGAATTCACAGAAACACCCAGAGCTTCGATGCGAGAACTCAAGTGGCCCTGT 763
 QY 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyIysSerLeuIysThrPro 201
 DB 764 CTCTCAGGCTCCCTGTCTCTCCCTGCACTGTCTTGCCTGTGGAGAGCCCTGAAGACCCC 823
 QY 202 ArgValValGlyGlyGluGluIaSerValAspSerTrpProTrpGlnValSerIleGln 221
 DB 824 CGTGTGTGGGTGGAGAGAGGCTCTGTGTGATTTGGCTTGGCAGGTCAAGATCCAG 883
 QY 222 TyrAspIysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAla 241
 DB 884 TACGACAAACAGACCTCTGTGAGAGGAGCATCTTGACCCCACTGGGTCTTCAACGGCA 943
 QY 242 AlaHisCysPheArgIysHisThrAspValPheAsnTrpIysValArgAlaGlySerAsp 261
 DB 944 GCCCATGTGTTACGAGAAACATACCGATGTGTTCACTGGAAGGTGGGCAAGCTCAAGC 1003
 QY 262 LysLeuGlySerPheProSerLeuAlaValAlaIysIleIleIleGluPheAsnPro 281
 DB 1004 AATCTGGGAGCTTCCCATCTCCGTGGCTGTGGCCAAAGATCATCATGAAATTCACACCC 1063
 QY 282 MetTyrProIysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
 DB 1064 ATGTACCCCAAGACATGATCATCGCCCTCAATGAAGCTGCACTTCCACTCTTCTCA 1123
 QY 302 GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro 321
 DB 1124 GGCACAGTCAAGCCCATCTGTCTGCTTCTTGTATGAGAGCTCACTCCAGCCACCCCA 1183
 QY 322 LeuTrpIleIleGlyTrpGlyPheThrIysGlnAsnGlyGlyLysMetSerAspIleLeu 341
 DB 1184 CTCTGATCATTTGATGGGCTTTTACGAAGCAGATGAGGAGAGATGTCTGACATACTG 1243

QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGln 361
 DB 1244 CTGACAGCGTCAGTTCACAGTTCATTTGACAGCACAGGTGCAATGACAGATCGTACAG 1303
 QY 362 GlyGluValThrGluLysMetCysAlaGlyIleProGluGlyValIleAspThrCys 381
 DB 1304 GGGAGAGTCAACCAAGAAAGTATGTGTGAGGATCCCGAAAGGGGTGTGACACCTGC 1363
 QY 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValIleGlyIle 401
 DB 1364 CAGGTGACAGTGTGGGCTCCCTGATGTACCATGTGACAGTGTGATGTGTGAGCTTC 1423
 QY 402 ValSerTrpGlyTyrGlyCysGlyIysProSerThrProGlyValTyrThrLysValSer 421
 DB 1424 GTTAGCTGGGCTATAGCTGTGCGGGGCGCGAGACCCAGAGATATACCAAGGTCTCA 1483
 QY 422 AlaTyrLeuAsnTrpIleTyrAsnValIleTrpIysAlaGluLeu 435
 DB 1484 GCTATCTCACTGGAATCTCAATGTCTGAAAGCTGAGCTG 1525

RESULT 11
 US-10-956-157-2292
 ; Sequence 2292, Application US/10956157
 ; Publication No. US20050118625A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wyeth
 ; APPLICANT: Mounts, William
 ; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
 ; FILE REFERENCE: 031896-043000 (AM 101081)
 ; CURRENT APPLICATION NUMBER: US/10/956,157
 ; NUMBER OF SEQ ID NOS: 319805
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 2292
 ; LENGTH: 2104
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-956-157-2292

Alignment Scores:
 Pred. No.: 1,566-282 Length: 2104
 Score: 2337.00 Matches: 434
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 99.79% Indels: 0
 DB: 9 Gaps: 0

US-10-803-530-2 (1-435) x US-10-956-157-2292 (1-2104)

QY 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValIysProLeuArgIysProArg 21
 DB 242 GATCTGACAGTGAATCACTCTGAAACAGCTCCGATGCAAAACCCCTGCGAAACCCCGT 301
 QY 22 IlePrometGluThrPheArgIysValGlyIleProIleIleIleAlaLeuLeuSerLeu 41
 DB 302 ATCCCAATGAGAGACCTTCAAGAAAGGTGGGATCCCATCATCATAGACATCACTGAGCTG 361
 QY 42 AlaserIleIleIleValIleValIleValIleValIleValIleValIleValIleValIle 61
 DB 362 GCGATATCATCATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 421
 QY 62 CysGlyGlnProLeuHisPheIleProArgIysGlnLeuCysAspGlyGluLeuAspCys 81
 DB 422 TGGGGAGAGCTTCCACTTTCATCCGAGAAAGCAGCTGTGTGACGAGAGCTGTGACTGT 481
 QY 82 ProLeuGlyGluAspGlnGluHisCysValIysSerPheProGluGlyProAlaValAla 101
 DB 482 CCCTTGGGGAGAGACGAGAGACCTGTGTCAAGAGCTTCCCGAAAGGCTGTGACAGTGGCA 541
 QY 102 ValArgLeuSerIysAspArgSerThrLeuGlnValIleuAspSerAlaThrGlyAsnTrp 121
 DB 542 GTCCGCTCTCCAAAGACCGATCCACATGACAGGTGTGTGACTGCGCCACAGGAACTGG 601

PRIOR APPLICATION NUMBER: US 60/315,287
PRIOR FILING DATE: 2001-08-27
PRIOR APPLICATION NUMBER: US 60/350,666
PRIOR FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/372,246
PRIOR FILING DATE: 2001-04-12
NUMBER OF SEQ ID NOS: 163
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 88
LENGTH: 1314
TYPE: DNA
ORGANISM: Homo sapiens
US-10-173-999-88

Alignment Scores:
Pred. No.: 7,58e-283 Length: 1314
Score: 2337.00 Matches: 434
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.79% Indels: 0
DB: Gaps: 0

US-10-803-530-2 (1-435) x US-10-173-999-88 (1-1314)

Qy 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValProLeuArgProArg 21
Db 10 GATCCTGACAGTGAACCTCTGAAAGCCTCGATGCAAACTTCGCAAACTTCCTG 69
Qy 22 IleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuSerLeu 41
Db 70 ATCCCATGAGACCTTCAGAAAGGTGGGATCCCATCATCATGCACTACTAGCCTG 129
Qy 42 AlaSerIleIleIleValIleValIleuIleuValIleuAspLysTrpPheLeu 61
Db 130 GCGAGTATCATCATTTGTGTGTCTCTCATCAAGTGATTCGAAATAATGACTTCTC 189
Qy 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys 81
Db 190 TGGCGGACAGCTCTCCACTTCATCCCGAAGAGCAGCTGTGTGACGGAGCTGACTGT 249
Qy 82 ProLeuGlyGluAspGlnGluHisCysValLysSerPheProGluGlyProAlaValAla 101
Db 250 CCTTTGGGGAGAGAGAGAGACGCTGTCAAGACTTCCCGAAGGGCTGACAGTGCA 309
Qy 102 ValArgLysSerLysAspArgSerThrLeuGlnValIleuAspSerAlaThrGlyAsnTrp 121
Db 310 GTCCGCTCTCCAAAGACCGATCCACACTGCGAGTGTGACTCGGCAACAGGAACTGG 369
Qy 122 PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
Db 370 TTCTGTGCTGTTTGGACACTTCACAGAGCTCTCGTGAGACAGCGCTGTAGGCAATG 429
Qy 142 GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
Db 430 GGCCTACAGAGCAAAACCACTTCAGAGCTGTGGAGATTGGCCAGACAGCATCTGAT 489
Qy 162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
Db 490 GTTGTGTAATACAGAAACAGCCAGAGCTTCGATGCGGAATCTCAAGTGGGCGCTGT 549
Qy 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuThrPro 201
Db 550 CTCTCAGGCTCCCTGCTGCTCCCTGCACTGTCTTGCTGTGGGAAGAAGCTGAAGACCCC 609
Qy 202 ArgValValGlyGlyGluGluAlaSerValAspSerTrpProGlnValSerIleGln 221
Db 610 CGTGTGGTGGTGGGAGAGGCGCTCTGTGATTTCTTGCGCTTGAGAGTCAAGATCCAG 669
Qy 222 TyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAla 241
Db 670 TACGACAAACAGACAGCTGTGTGAGGAGACATCTGGACCCCACTGGGTCTCTCAGGGA 729
Qy 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpValArgAlaGlySerAsp 261

Db 730 GCCCAGCTCTTCAGAAACATACGATGTGTTCAACTGGAAGGTGGCGGACGCTTCAGAC 789
Qy 262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleIleGluPheAsnPro 281
Db 790 AAATGGGACAGCTTCCATCCCTGGCTGTGGCAAGATCATCATCATTTGAATCAACCCC 849
Qy 282 MetTrpProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db 850 ATGTACCCCAAAAGACATGACATGACCTCATGAAAGCTGCAAGTCCCATCATCTTCTCA 909
Qy 302 GlyThrValArgProIleCysLeuProPhePheAspGluLeuThrProAlaThrPro 321
Db 910 GGCACAGTACAGGCCCATCTGCTCTCTCTTGTATGAGAGCTCATCTCCAGCACCCCA 969
Qy 322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu 341
Db 970 CTCTGTATCATTTGATGTGGGCTTTTACAAAGAGATGAGGAAAGATGTGACATACTG 1029
Qy 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTrpGln 361
Db 1030 CTGACGGGTGACGTCAGGTGATTTGACAGACACAGGTGCAATGACAGATGCTTACAG 1089
Qy 362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys 381
Db 1090 GGGGAAGTCAACCGAAGATGATGTGTGACAGCATCCCGAAGGGGTGTGACACTGTC 1149
Qy 382 GlnGlyAspSerGlyGlyProLeuMetTrpGlnSerAspGlnTrpHisValValGlyIle 401
Db 1150 CAGGTGACAGTGGTGGGCCCCCTGATGTACCAATCTGACAGTGGATGTGTGGGATC 1209
Qy 402 ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSer 421
Db 1210 GTTAGCTGGGGCTATGTGTGTGGGGGCCCGAGCACCCCAAGATATACCAAGGTCTCA 1269
Qy 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db 1270 GCTTATCTCACTGATCTACAAATGTCTGGAAGGCTGAGCTG 1311

RESULT 10
US-09-851-588-7
Sequence 7, Application US/09851588
Patent No. US20020042067A1
GENERAL INFORMATION:
APPLICANT: Mack, David
APPLICANT: Gish, Kurt C.
APPLICANT: Wilson, Keith E.
TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSING COLORECTAL CANCER, COMPOSITIONS, AND
FILE REFERENCE: A-68829-1/DJB/JCD/AMS
CURRENT APPLICATION NUMBER: US/09/851,588
CURRENT FILING DATE: 2001-09-24
PRIOR APPLICATION NUMBER: US 09/642,252
PRIOR FILING DATE: 2000-08-17
PRIOR APPLICATION NUMBER: US 09/656,002
PRIOR FILING DATE: 2000-09-06
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn version 3.1
SEQ ID NO 7
LENGTH: 2081
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (215)..(1528)
OTHER INFORMATION:
US-09-851-588-7

Alignment Scores:
Pred. No.: 1,54e-282 Length: 2081
Score: 2337.00 Matches: 434
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

```
FILE REFERENCE: 018501-012500US
CURRENT APPLICATION NUMBER: US/10/295,027
CURRENT FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: US 09/663,733
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: US 60/350,666
PRIOR FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/335,394
PRIOR FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: US 60/332,464
PRIOR FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: US 60/334,393
PRIOR FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: US 60/340,376
PRIOR FILING DATE: 2001-12-14
PRIOR APPLICATION NUMBER: US 60/347,211
PRIOR FILING DATE: 2002-01-08
PRIOR APPLICATION NUMBER: US 60/347,349
PRIOR FILING DATE: 2002-01-10
PRIOR APPLICATION NUMBER: US 60/355,250
PRIOR FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: US 60/356,714
PRIOR FILING DATE: 2002-02-13
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1386
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 979
LENGTH: 1314
TYPE: DNA
ORGANISM: Homo sapiens
US-10-295-027-979

Alignment Scores:
Pred. No.: 7,58e-283 Length: 1314
Score: 2337.00 Matches: 434
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.79% Indels: 0
Gaps: 0

US-10-803-530-2 (1-435) x US-10-295-027-979 (1-1314)

QY 2 AspProAspSerArgGlnProLeuAsnSerLeuAspValIysProLeuArgLysProArg 21
Db 10 GATCTCAACATGATCACTCACTGCAAGCCTCGATTCGAACCCCTGGCAACCCCTG 69
QY 22 IleProMetGlnThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSerLeu 41
Db 70 ATCCCATGAGAGCTTCAGAAAGGTGGGATCCCATCATCATAGCAGCTAGAGCTG 129
QY 42 AlaSerIleIleIleValIleValIleValIleValIleValIleValIleValIleVal 61
Db 130 GCGAGTATCATCTGTTGGTTGTTCTTCATCAAGTATTCGTGATTAATCACTTCTTC 189
QY 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGlnLeuAspCys 81
Db 190 TGGGGAGAGGCTCTCCACCTTCATCCGAGGAGCACTGTGTGACGAGAGCTGACTGT 249
QY 82 ProLeuGlyGlnAspGlnGlnHisCysValIysSerPheProGlnGlyProAlaValAla 101
Db 250 CCGTGGGGAGAGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 309
QY 102 ValArgLeuSerIysAspArgSerThrGlnGlnValIleAspSerAlaThrGlyAsnTrp 121
Db 310 GTCCGCTCTCCAGAGAGGAGTCCACATGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 369
QY 122 PheSerAlaCysPheAspAsnPheThrGlnAlaLeuAlaGlnThrAlaCysArgGlnMet 141
Db 370 TTCTGCTGCTGTTTGCAGCACTTCACAGAGCTCTGCTGAGACAGCCTGTAGGAGATG 429
QY 142 GlyTyrSerSerIysProThrPheArgAlaValGlnIleGlyProAspGlnAspLeuAsp 161
Db 430 GGCTACAGAGCAAAACCCACTTTCAGAGCTGTGAGATTTGGCCGAGACAGAGATCTGAT 489
```

```
QY 162 ValValGlnIleThrGlnAsnSerGlnGlnLeuAsnMetArgAsnSerSerGlyProCys 181
Db 490 GTTGTGAATTCACAGAAACAGCAGAGCTTCGATCGGAGCTCAAGTGGGCTGTG 549
QY 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuIysThrPro 201
Db 550 CTCTCAGGCTCCCTGATCTCCCTCGCAGCTGTCTTGGCTGGGAGAGGCTGAAAGCCCC 609
QY 202 ArgValValGlyGlyGlnGlnAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
Db 610 CGTGTGTGGTGGGAGAGAGGCTTGTGATTTCTTGACCTTGGCAGGTCAAGATCCAG 669
QY 222 TyrAspLysGlnHisValCysGlyIysSerIleLeuAspProHisTrpValLeuThrAla 241
Db 670 TACGACAAACAGACAGCTGTGTGAGGAGCACTCGAGACCCCATGTGGCTTCACGGCA 729
QY 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
Db 730 GCCCACTGCTTACAGAAACATACCATGATGTTCACTGGAAGTGGGAGGCTCAGAC 789
QY 262 LysLeuGlySerPheProSerLeuAlaValAlaValIleIleIleGlnPheAsnPro 281
Db 790 AAACGTGGGAGCTTCCATCCCTGCTGTGGCAGATCATCATCATTTGAATTCAAACCC 849
QY 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db 850 ATGTACCCCAAGACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 321
QY 302 GlyThrValArgProIleCysLeuProPhePheAspGlnGlnLeuThrProAlaThrPro 341
Db 910 GGCACAGTACGGCCATCTGTCTGCTCTTCTTGTATGAGAGTCACTCCAGCCACCCCA 361
QY 322 LeuTrpIleIleGlyTrpGlyPheThrIleGlnGlnGlnGlyLysMetSerAspIleLeu 381
Db 970 CTCTGATCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 401
QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGln 421
Db 1030 CTGACAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 441
QY 362 GlyGlnValIleThrGlnLysMetMetCysAlaGlyIleProGlnGlyGlyValAspThrCys 461
Db 1090 GGGAGAGTCAACGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 481
QY 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValIleGlyIle 501
Db 1150 CAGGCTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 521
QY 402 ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValIleTyrThrLysValSer 541
Db 1210 GTTAGCTGGGGCTATGCTGCTGGGGGGCCGAGAGACCCCAAGATTAACCAAGATGCTCA 561
QY 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGlnLeu 581
Db 1270 GCTTATCTCAACTGATCTCAATATGCTGGAAGGCTGAGCTG 601

RESULT 9
US-10-173-999-88
Sequence 88, Application US/10173999
Publication No. US20040005563A1
GENERAL INFORMATION:
APPLICANT: Mack, David H.
APPLICANT: Gish, Kurt C.
APPLICANT: Eos Biotechnology, Inc.
TITLE OF INVENTION: Methods of Diagnosis of Ovarian Cancer, Compositions
TITLE OF INVENTION: and Methods of Screening for Modulators of Ovarian
FILE REFERENCE: 018501-002420US
CURRENT APPLICATION NUMBER: US/10/173,999
PRIOR APPLICATION NUMBER: US 60/299,234
PRIOR FILING DATE: 2001-06-18
```

CURRENT APPLICATION NUMBER: US/10/295,027
CURRENT FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: US 09/663,733
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: US 60/350,666
PRIOR FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/335,394
PRIOR FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: US 60/332,464
PRIOR FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: US 60/334,393
PRIOR FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: US 60/340,376
PRIOR FILING DATE: 2001-12-14
PRIOR APPLICATION NUMBER: US 60/347,211
PRIOR FILING DATE: 2002-01-08
PRIOR APPLICATION NUMBER: US 60/347,349
PRIOR FILING DATE: 2002-01-10
PRIOR APPLICATION NUMBER: US 60/355,250
PRIOR FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: US 60/356,714
PRIOR FILING DATE: 2002-02-13
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1386
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 830
LENGTH: 1314
TYPE: DNA
ORGANISM: Homo sapiens
US-10-295-027-830

Alignment Scores:

Pred. No.: 7,58e-283 Length: 1314
Score: 2337.00 Matches: 434
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.79% Indels: 0
DB: 6 Gaps: 0

US-10-803-530-2 (1-435) x US-10-295-027-830 (1-1314)

QY 2 AspProAspSerAspGlnProLeuAnSerLeuAspValLysProLeuArgLysProArg 21
Db 10 GATCCTGACAGTGAACCTCTGAAACAGCCTCGATGTCAAACCCCTGGCGAAACCCCGT 69
QY 22 IleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuSerLeu 41
Db 70 ATCCCATGAGACCTTCAGAAAGGTGGGATCCCATCATCATGACACTACTGAGCCTG 129
QY 42 AlaSerIleIleIleValValValLeuIleLysValIleLeuAspLysTyrThrPheLeu 61
Db 130 GCGATATCATCATGTGGTGTCTCTCATCAAGGATTCGATTAATAACTACTCTCCTC 189
QY 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys 81
Db 190 TCGGGGAGCCCTCTCATCTCATCCGAGGAAGACGTGTGAGAGGAGGAGTGAAGTGT 249
QY 82 ProLeuGlyGluAspGluGluHisCysValLysSerPheProGluGlyProAlaValAla 101
Db 250 CCCTTGGGGAGAGAGAGAGAGACGTGTCAAGAGCTTCCCGAAGGGCGCTGCACTGGCA 309
QY 102 ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlySerTrp 121
Db 310 GTCCCTCTCTCCAAAGACCGATCCACACTGAGAGGTCTGACTCGGCCACAGGAACTGG 369
QY 122 PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
Db 370 TTCTCTGCTGTTTGGACAACTTCAAGAAAGCTCTGCTGAGACAGCCTGTAGGGAGAGT 429
QY 142 GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
Db 430 GGCCTACAGAGCAAAACCCACTTTCAGAGCTGTGGAGATTGGCCAGACAGAGATCTGAT 489

QY 162 ValValGluIleThrGluAnSerGlnGluLeuArgMetArgAnSerSerGlyProCys 181
Db 490 GTTGTGTAATTCACAGAAACAGCCAGAGCTTGCATGCGGAATCAAGTGGCCCTGT 549
QY 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro 201
Db 550 CTCACAGCTCCCTGGTCTCTCTGACACTGCTTCTCTGTGGAAAGCCTTAAGACCCCC 609
QY 202 ArgValValGlyGlyGluGluAlaSerValAspSerTrpProTyrGlnValSerIleGln 221
Db 610 CTTGTGTGGGTGGGAGAGAGGCTCTGTGATTTCTTGAGCTTGGCAGCTAGCATCCAG 669
QY 222 TyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAla 241
Db 670 TAGCAAAACAGCAGCTGTGTGAGGAGCATCTTGAGCCCACTGGGTCTTCACGGCA 729
QY 242 AlaHisCysPheArgLysIleThrAspValPheAsnTrpLysValAlaGluAlaGlySerAsp 261
Db 730 GCCCAGCTGCTTCAGAGAAACATACCGATGTGTTCAACTGGAAGGTGGCGGAGCTCAGAC 789
QY 262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsnPro 281
Db 790 AACTGGGCACTTCCCATCTCTGCTGTGGCCAAAGATCATCATTTGAATTCACACCC 849
QY 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db 850 ATGTACCCCAAAAGACATGATGATCGCCCTCATGAAGTGCAGTTCCCATCATTCTCA 909
QY 302 GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro 321
Db 910 GGCACAGTCAGGCCCATCTGTCTGCCCTTCTTATGAGGAGCTCACCTCCAGCCACCCCA 969
QY 322 LeuTrpIleIleGlyTyrPglPheThrLysGlnAnGlyLysMetSerAspIleLeu 341
Db 970 CTGTGATCATTTGATGGGCTTTTACAGAGAAATGAGGAAAGATGTCTGACATCTTG 1029
QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGln 361
Db 1030 CTGCAGGGCTGATGTCAGGTGATTCATTCAGCAGCAGCGGTGCATGACAGAGTGCATCCAG 1089
QY 362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyValAspThrCys 381
Db 1090 GGGGAGTCACCCGAGAGATATGTGTCAGGCAATCCCGAAGGGGTGTGACACCTGC 1149
QY 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle 401
Db 1150 CAGGTGACAGTGTGGGCTCTGATGTACAAATCTGACAGTGCAGTGTGTGGGCACTC 1209
QY 402 ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSer 421
Db 1210 GTTAGCTGGGCTATGCTGTGGGGGCCGAGCACCCAGAGATATACCAAGAGTCTCA 1269
QY 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db 1270 GCTTATCTCACTGATCTCAATGTCTGAAAGGCTGAGCTG 1311

RESULT 8
US-10-295-027-979
Sequence 979, Application US/10295027
Publication No. US2003023250A1
GENERAL INFORMATION:
APPLICANT: Afar, Daniel
APPLICANT: Aziz, Natasha
APPLICANT: Ginsberg, Wendy M.
APPLICANT: Glen, Kurt C.
APPLICANT: Glynn, Richard
APPLICANT: Hevezi, Peter A.
APPLICANT: Mack, David H.
APPLICANT: Murray, Richard
APPLICANT: Watson, Susan R.
APPLICANT: Bos Biotechnology, Inc.
TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
TITLE OF INVENTION: Methods of Screening for Modulators of Cancer

```

; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO: 790
; LENGTH: 1314
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-295-027-790

Alignment Scores:
Pred. No.: 7,58e-283      Length: 1314
Score: 2337.00           Matches: 434
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 99.79%             Indels: 0
DB: 6                       Gaps: 0

US-10-803-530-2 (1-435) x US-10-295-027-790 (1-1314)

QY 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValIleuAspProArg 21
DB 10 GATCTCAAGATGATCAACCTCTGAACAGCTCGATGTCAACCCCTGGCAACCCCT 69
QY 22 IleProMetGlnThrPheArgIleValGlyIleProIleIleIleLeuSerLeu 41
DB 70 ATCCCAATGAGACCTTCAGAAAGGTGGGATCCCATCATCATACACTGAGCCCTG 129
QY 42 AlaSerIleIleIleValIleValIleuIleValIleuAspIleuIleuSerLeu 61
DB 130 GCGAGTATCATGATGTGCTGCTCCATCAAGGATGATTCGTGATTAATTAATCTTCTC 189
QY 62 CysGlyGlnProLeuHisPheIleProArgIleGlnLeuCysAspGlyIleuAspCys 81
DB 190 TGGGGGACGCTCTCCACTTCATCCCGAAGAGACGCTGTGACGAGAGCTGACTGT 249
QY 82 ProLeuGlyValAspGlnGlnHisCysValIleuSerPheProGlnGlyProAlaValAla 101
DB 250 CCTTGGGGAGAGACAGAGAGACCTGTGTCAAGAGCTTCCCGAAGGCGCTGCACTGCA 309
QY 102 ValArgLeuSerIleAspArgSerThrLeuGlnValIleuAspSerIleThrGlyAsnTrp 121
DB 310 GTCCGCTCTCCAAAGACCGATCCACATCGCAGGTGCTGACTCGGCCACAGGGAACTGG 369
QY 122 PheSerAlaCysPheAspAsnThrThrAlaLeuAlaGluThrAlaCysArgGlnMet 141
DB 370 TTCTCTGCTGTTTGGCAAACTTCAAGAGCTCTCGCTGAGACAGCTGTAGGCAATG 429
QY 142 GlyTyrSerSerIleProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
DB 430 GGCTACAGACGAAACCACTTCAGACCTGTGAGATGGCCACAGACAGAGATCTGGAT 489
QY 162 ValValGluIleThrGluAsnSerGlnIleuAsnMetArgAsnSerSerGlyProCys 181
```

```

DB 490 GTTGTGAATTCACAGAAAACAGCCAGAGCTTCGATCGGAGACTCAAGTGGCCCTGT 549
QY 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyIleuSerLeuIleuThrPro 201
DB 550 CTCTCAGGCTCCTGTGTTCTCCCTGCACTGTCTTGGCTGTGGAGAGGCTCGAAGACCCCC 609
QY 202 ArgValAlaGlyIleGlnIleuAspSerValAspSerTrpProTrpGlnValSerIleGln 221
DB 610 CGTGTGGTGGTGGGAGAGAGGCTGTGATTTCTTGACCTTGGCAGGTCAAGATCCAG 669
QY 222 TyrAspIleGlnHisIleValCysGlyIleuSerIleuAspProHisIleuThrAla 241
DB 670 TACGACAAACAGACGCTGTGTGAGAGAGACATCTGCAACCCCACTGGGTCTTCAAGCA 729
QY 242 AlaHisCysPheArgIleHisThrAspValPheAsnTrpValArgAlaGlySerAsp 261
DB 730 GCCCACTGCTTCAGGAACATACCATGTGTTCACTGAAAGTGGGGGAGGCTCAGAC 789
QY 262 IysLeuGlySerPheProSerLeuAlaValAlaIleuIleIleIleGlnPheAsnPro 281
DB 790 AAACCTGGGAGCTTCCCATCTCCGTGGCTGGCAAGATCATCATTTGATTTCAACCCC 849
QY 282 MetTyrProIysAspAsnAspIleAlaLeuMetIleuGlnPheProLeuThrPheSer 301
DB 850 ATGTACCCCAAGACATGACATGACCTCTCATGAGGCTGCACTTCCACTCATCTTCTCA 909
QY 302 GlyThrValArgProIleCysLeuProPhePheAspGlnIleuThrProAlaThrPro 321
DB 910 GGCACAGTCAGGCCCATGTGTCTGCCCTTTTATAGAGCTCATCTCCAGCCACCCCA 969
QY 332 LeuTrpIleIleGlyTrpGlyPheThrIleGlnAsnGlyIleuMetSerAspIleLeu 341
DB 970 CTCTGATCATTTGATGAGGGCTTTTACAGACAGATGAGGAGGATGTGACATACCTG 1029
QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaIleuGln 361
DB 1030 CTGACAGGCTGATAGTCCAGGCTGATTTGACAGACACAGGAGCAATCAGACGATGGTACCAG 1089
QY 362 GlyGluValThrGlnIleuMetCysAlaGlyIleProGlnGlyIleValAspThrCys 381
DB 1090 GGGGAGATCACCGAGAGATGATGTGACAGCATCCCGAAGGGGATGTGGACACTGC 1149
QY 382 GlnIleuAspSerGlyIleProLeuMetTyrGlnSerAspGlnTrpHisValAlaGlyIle 401
DB 1150 CAGGCTACAGTGTGGTGGCCCTGATGTACCAATCTGACACAGTGGCATGTGGGCACTC 1209
QY 402 ValSerTrpGlyTyrGlyCysGlyIleProSerThrProGlyValIleuThrIleuValSer 421
DB 1210 GTTAGCTGGGGCTATGTGCTGCGGGGGCCGAGACACCCCAAGAGTATACCAAGGTCTCA 1269
QY 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpIleAlaGluLeu 435
DB 1270 GCCTATCTCAATCGATCTTACATGTCTGGAAGGCTGACTG 1311

RESULT 7
US-10-295-027-830
; Sequence 830, Application US/10295027
; Publication No. US2003023350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natsheh
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
```

```
/ PRIOR APPLICATION NUMBER: US 09/663,733
/ PRIOR FILING DATE: 2000-09-15
/ PRIOR APPLICATION NUMBER: US 60/350,666
/ PRIOR FILING DATE: 2001-11-13
/ PRIOR APPLICATION NUMBER: US 60/335,394
/ PRIOR FILING DATE: 2001-11-15
/ PRIOR APPLICATION NUMBER: US 60/332,464
/ PRIOR FILING DATE: 2001-11-21
/ PRIOR APPLICATION NUMBER: US 60/334,393
/ PRIOR FILING DATE: 2001-11-29
/ PRIOR APPLICATION NUMBER: US 60/340,376
/ PRIOR FILING DATE: 2001-12-14
/ PRIOR APPLICATION NUMBER: US 60/347,211
/ PRIOR FILING DATE: 2002-01-08
/ PRIOR APPLICATION NUMBER: US 60/347,349
/ PRIOR FILING DATE: 2002-01-10
/ PRIOR APPLICATION NUMBER: US 60/355,250
/ PRIOR FILING DATE: 2002-02-08
/ PRIOR APPLICATION NUMBER: US 60/356,714
/ PRIOR FILING DATE: 2002-02-13
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 1386
/ SOFTWARE: Patent In Ver. 2.1
/ SEQ ID NO 778
/ LENGTH: 1314
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-295-027-778

Alignment Scores:
Pred. No.: 7,586-283 Length: 1314
Score: 2337.00 Matches: 434
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.79% Indels: 0
DB: 6 Gaps: 0

US-10-803-530-2 (1-435) x US-10-295-027-778 (1-1314)
QY 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValIysProLeuArgIysProArg 21
DB 10 GATCTGTGACGTGATCAACCTCTGAAACAGCTCGATGTCMAACCCCTGGCAAAACCCCGT 69
QY 22 IleProMetGlnThrPheArgIysValGlyIleProIleIleIleIleValLeuLeuSerLeu 41
DB 70 ATCCCACTGAGAACCTTTCAGAAAGGTGGGAGATCCCATCATCATATGACACTAGAGCCGTG 129
QY 42 AlaSerIleIleIleValIleValIleValIleValIleValIleValIleValIleValIle 61
DB 130 GCGAGTATCATCATGTTGTTGTTCTCTCATCAAGGTGATTCGATTAATACTACTTCTCTC 189
QY 62 CysGlyGlnProLeuHisPheIleProArgIysGlnLeuCysAspGlyGlnLeuAspCys 81
DB 190 TCGGGGAGCCCTCTCCACTTCATCCCGAGGAGAGCTGTGTGACGGAGAGCTGAGCTGT 249
QY 82 ProLeuGlyGlnAspGlyGlnIleCysValIlysSerPheProGlyGlyProAlaValAla 101
DB 250 CCTTGGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 309
QY 102 ValArgLeuSerIysAspArgSerThrLeuGlnValIleuAspSerAlaThrGlyAsnTyr 121
DB 310 GTCCGCTCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 369
QY 122 PheSerAlaCysPheAspAsnSerThrGlnAlaValIleValIleValIleValIleValIle 141
DB 370 TTCTGTGCTGTTTGCAGAACTTCACAGAAAGCTTCGCGAGAGAGAGAGAGAGAGAGAGAGAG 429
QY 142 GlyTyrSerSerIysProThrPheArgAlaValIleGlyProAspGlnAspLeuAsp 161
DB 430 GGCTACAGAGCAAAACCACTTTCAGAGCTGTGGAGATGGGCCAGAGAGAGATCTGAGAT 489
QY 162 ValValIleGlnIleThrGlnAsnSerGlnIleuAspGlyMetArgAsnSerSerGlyProCys 181
```

```
DB 490 GTTGTGAATTCACAGAAAACAGCAGAGCTTCGCAATGGGAACTCAAGTGGCCCTGT 549
QY 182 LeuSerGlySerLeuValIleSerLeuHisCysLeuAlaCysGlyIysSerLeuIysThrPro 201
DB 550 CTCACAGCTCCCTGGGTCTCTCCCTGCACTGTCTTCCCTGTGGAGAGAGCTCGAAGACCCC 609
QY 202 ArgValIleGlyGlyGlnIleValIleAspSerThrProTyrGlnValIleGln 221
DB 610 CGTGTGGGTGGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 669
QY 222 TyrAspIysGlnIleValIleCysGlyIysSerIleLeuAspProHisThrValIleThrAla 241
DB 670 TACACAAACAGCAGCTCTGTGGAGAGAGATCTTGACCCCACTGGGTCTTCACAGCA 729
QY 242 AlaHisCysPheArgIysHisIleThrAspValPheAsnThrIysValIleAspAlaGlySerAsp 261
DB 730 GCCCAGCTGTTACAGAAACATACCGATGTTCACATGAAAGGTGGCGAGGCTGACAG 789
QY 262 LysLeuGlySerPheProSerLeuAlaValAlaValIleIleIleIleGlnPheAsnPro 281
DB 790 AAATCGGGAGCTTCCATCCCTGTGTGGCCAAAGATCATCATCATTTGAATTCACCCC 849
QY 282 MetTyrProIysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
DB 850 ATGTACCCCAAAAGACATGACATGCCCTCATGAAGCTGACGTTCCCACTCATTTCTCA 909
QY 302 GlyThrValArgProIleCysLeuProPheAspGlnIleuThrProAlaThrPro 321
DB 910 GGCACAGTCAGGCCCATCTGTCTCCCTCTTTATGAGAGAGTCATCCACGCCACCCCA 969
QY 322 LeuThrIleIleGlyTyrGlyPheThrIleGlnAsnGlyIysMetSerAspIleLeu 341
DB 970 CTCTGATCATTTGATGGGGCTTTTACAGAGCATGAGAGAGAGAGATGTCTGACATCTG 1029
QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGln 361
DB 1030 CTGCAGGCTGATGCTCAGATTCATTCAGACAGCAGAGTGCATATGACAGAGAGCTGACAG 1089
QY 362 GlyIleValIleThrGlnIysMetMetCysAlaGlyIleProGlyGlyIleAspThrCys 381
DB 1090 GGGAGAGTACACGAGAGAGATGATGTGACAGCATCCCGAAGGGGGTGGACACCTGC 1149
QY 382 GlnGlyAspSerGlyIysProLeuMetTyrGlnSerAspGlnThrIleValIleGlyIle 401
DB 1150 CAGGTGACAGTGTGGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1209
QY 402 ValSerTyrGlyIysGlyIysCysGlyIysProSerThrProGlyValIleThrIleValSer 421
DB 1210 GTTAGCTGGGGCTATGGCTGGGGGGCCCGAGCACCCCAAGAGATATACCAAGAGTCTCA 1269
QY 422 AlaTyrLeuAsnThrIleTyrAsnValIleIleValIleValIleValIleValIleValIle 435
DB 1270 GCCTATCTCAACTGATTCATCATATGTCTGAAAGGCTGAGCTG 1311

RESULT 6
US-10-295-027-790
/ Sequence 790, Application US/10295027
/ Publication No. US2003023250A1
/ GENERAL INFORMATION:
/ APPLICANT: Afar, Daniel
/ APPLICANT: Aziz, Natasha
/ APPLICANT: Ginsberg, Wendy M.
/ APPLICANT: Gish, Kurt C.
/ APPLICANT: Glynn, Richard
/ APPLICANT: Hevezl, Peter A.
/ APPLICANT: Mack, David H.
/ APPLICANT: Murray, Richard
/ APPLICANT: Watson, Susan R.
/ APPLICANT: Eos Biotechnology, Inc.
/ TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
/ FILE REFERENCE: 018501-012500US
/ CURRENT APPLICATION NUMBER: US/10/295, 027
```

```

; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 133
; LENGTH: 1314
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-295-027-133

Alignment Scores:
Pred. No.: 7,58e-283 Length: 1314
Score: 2337.00 Matches: 434
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.79% Indels: 0
DB: 6 Gaps: 0

US-10-803-530-2 (1-435) x US-10-295-027-133 (1-1314)
QY 2 AspProAspSerAspGlnProLeuAenSerLeuAspValLysProLeuArgLysProArg 21
Db 10 GATCCTGACGATGATCACTCTGACAGCCTGAGTGTCAAAACCCCTGGCGAAACCCCTGT 69
QY 22 IleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSerLeu 41
Db 70 ATCCCATGAGAGCCTTCAGAAAGTGGGGATCCCATCATCATATACACTGAGACCTG 129
QY 42 AlaSerIleIleIleIleValIleValIleValIleValIleValIleValIleValIle 61
Db 130 GCGAGTATCATCATGTGTGTGTCTCATCAAGTGTATTCGTGATTAATTAATTAATTCCT 189
QY 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys 81
Db 190 TGGGGGACAGCTTCCACTTTCATCCGAGAGAGCAGCTGTGTGACGAGAGCTGACCTGT 249
QY 82 ProLeuGlyGluAspGluGluHisCysValLysSerPheProGluGlyProAlaValAla 101
Db 250 CCCTTGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 309
QY 102 ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp 121
Db 310 GTCCGCTCTCCAGAGAGAGATCCACATGACAGGTGCTGAGCTCGGCCACAGGGAGACTGG 369
QY 122 PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
Db 370 TTCTCTGCTGTGTTGACAACTTTCACAGAGCTCTCGCTGAGACAGCTGTGTGGCAGATG 429
QY 142 GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGluAspLeuAsp 161
Db 430 GGCCTACAGAGCAAAACCCACTTTCAGAGCTGTGGAGATTTGGCCCAACAGAGATCTGAT 489
QY 162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
Db 490 GTTGTGAATCACAGAAACAGAGAGAGCTTCCCATGCGGAATCTCAAGTGGGCTCTGT 549
```

```

QY 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuIysThrPro 201
Db 550 CTCTCAGAGCTCCCTGTCTCCCTGACAGCTGTCTGTGCTGTGGAGAGAGCTGAAAGCCCC 609
QY 202 ArgValIleGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
Db 610 CGGTGTGTGGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 669
QY 222 TyrAspLysGlnHisValCysGlyIleSerIleLeuAspProHisTrpValLeuThrAla 241
Db 670 TACGACAAACAGACAGCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 729
QY 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
Db 730 GCCCACTGCTTCAGAAACATACCATGATGTTCACTGAAAGTGTGGGAGGCTCAGAC 789
QY 262 LysLeuGlySerPheProSerLeuAlaValAlaValIleIleIleGlnPheAsnPro 281
Db 790 AAACCTGGGAGCTTCCCATCCCTGGCTGTGGCAAGATCATCATTTGAATTCAAACCCC 849
QY 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db 850 ATGTACCCCAAGACAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 909
QY 302 GlyThrValArgProIleCysLeuProPheAspGluGluLeuThrProAlaThrPro 321
Db 910 GGCACAGTCAGGCCCATGTCTGTCTCTTGTATGAGAGAGTCACTCCAGCCACCCCA 969
QY 322 LeuTrpIleIleGlyTyrGlyPheThrIleGlnGlnGlyValLysMetSerAspIleLeu 341
Db 970 CTCTGATCATTTGATGAGGAGCTTTTACAGAGAGATGAGAGAGAGAGAGATGTGACATACGT 1029
QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGln 361
Db 1030 CTGACAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1089
QY 362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyValAspThrCys 381
Db 1090 GGGGAGTCAACCGAAGAGATGATGTGACAGGATCCCGGAGGGGTGTGGACACTGC 1149
QY 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValIleGlyIle 401
Db 1150 CAGGTGACATGTGTGAGGCTTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1209
QY 402 ValSerTrpGlyTyrGlyCysGlyIleProSerThrProGlyValTyrThrLysValSer 421
Db 1210 GTTAGCTGGGGCTATGCTGCTCGGGGGCCGAGACACCCACAGAGTATACACCAAGGTCTCA 1269
QY 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db 1270 GCTATCTCAACTGATCTACATATGTCTGGAAGGCTGAGCTG 1311

RESULT 5
US-10-295-027-778
; Sequence 778, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezl, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of diagnosis of cancer. Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/11/045,577
FILING DATE: 27-Jan-2005
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/659,151
FILING DATE: 11-Sep-2000
APPLICATION NUMBER: 09/008,271
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Mohan-Peterson, Sheela
REGISTRATION NUMBER: 41,201
REFERENCE/DOCKET NUMBER: PF-0458 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 2038 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: COLNOT13
CLONE: 1337018
SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-11-045-577-18

Alignment Scores:
Pred. No.: 1,11e-282 Length: 2038
Score: 2338.00 Matches: 434
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.77% Mismatches: 0
Query Match: 99.83% Indels: 0
DB: 10 Gaps: 0

US-10-803-530-2 (1-435) x US-11-045-577-18 (1-2038)

QY 1 MetAspProAspSerAspGlnProLeuAsnSerLeuAspValysProLeuArglyAsPro 20
DB 200 ATGGATCCTGACAGTCAACCTCTGAAACAGCTCGATGTCAACCCCTGGGCAACCC 259
QY 21 ArgIlePrometGluThrPheArglyValGlyIleProIleIleIleAlaLeuLeuSer 40
DB 260 CGTATCCCATCGAAGACCTTCGAAAGGATGGGATCCCATCATCATGACATCTGAGC 319
QY 41 LeuAlaSerIleIleValValValLeuIleValIleValIleLeuAspLysTyrPhe 60
DB 320 CTGGGAGATCATCATGTGTGTGTCTCATCAAGGATTCGTGATTAATTAATTAATTC 379
QY 61 LeuCyGlyGlnProLeuHisPheIleProArglyGlnLeuCyAspGlyGluLeuAsp 80
DB 380 CTCTGGGGGCAAGCTCTCCACTTCATCCGAGGAGAGCTGTGTGACGAGAGCTGAC 439
QY 81 CysProLeuGlyGluAspGluGlnHisCysValysSerPheProGluGlyProAlaVal 100
DB 440 TGTCTCTGGGGGAG 499
QY 101 AlaValAlaGluSerLeuAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsn 120
DB 500 GCAGTCGCGCTCTCCAAAGGACGATCCACATGCGAGGCTGAGCTCGGCCACAGGAGAC 559
QY 121 TrpPheSerAlaCysPheAspAsnSerThrGluAlaLeuAlaGluThrAlaCysArgGln 140
DB 560 TGTCTCTCTGCTCTGTTTGCACAACTTCACAGAAAGCTCTGCTGAGAACCTGTAGGAG 619
QY 141 MetGlyTyrSerSerLeuProThrPheArgAlaValGlyIleGlyProAspGlnAspLeu 160
DB 620 ATGGCTACAGAGAAACCCACTTTCAGAGCTGTGAGATGGGCCACAGAGATCTG 679
QY 161 AspValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyPro 180
DB 680 GATGTTGTTGAATCAGAGAAACAGCCGAGAGCTTCGATCGGAGATCAAGTGGGCC 739

QY 181 CysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThr 200
DB 740 TGTCTCTCAGGCTCCCTGCTCTCCCTGCTCAGCTCTCTGCTGAGGAGAGAGCTGAGAGC 799
QY 201 ProArgValValGlyGlyGluGlnAlaSerValAspSerTrpProTrpGlnValSerIle 220
DB 800 CCCGCTGTGTGGGTGGGAG 859
QY 221 GlnTyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThr 240
DB 860 CAGTACACAAACAGCAGCTGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 919
QY 241 AlaAlaHisCysPheArglyShiSerThrAspValPheAsnTrpLysValArgAlaGlySer 260
DB 920 GCAGCCCACTGCTCTCAGAGAAACATACGATGTGTTCAACTGAGAGAGAGAGAGAGAG 979
QY 261 AspLysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsn 280
DB 980 GACAACTGGGCACTTCCCATCCCTGCTGTGTGCGCAAGATCATCATCATTAATTCAC 1039
QY 281 PrometTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPhe 300
DB 1040 CCCATGTACCCCAAGACAAAGATGATGCGCTCATVGAAGCTGCACTCCACTCATCTTC 1099
QY 301 SerGlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThr 320
DB 1100 TCAGGACAGTCAAGGCCCATGTGTCTGCTCTCTTTATGAGAGAGCTCATCTCCAGCCAC 1159
QY 321 ProLeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyLysMetSerAspIle 340
DB 1160 CCATCTGAGATCATTTGATGGGCTTTACAGACAGATGAGAGAGAGAGATGTCAGACTA 1219
QY 341 LeuLeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyr 360
DB 1220 CTGCTGACAGGCGTCAAGTCCAGATGATGACAGACACAGGTCATGACAGAGAGCGTAC 1279
QY 361 GlnGlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThr 380
DB 1280 CAGGGAGAGTCAACCGAGAAATGATGTGTGACAGCATCCCGAGAGAGAGAGAGAGAG 1339
QY 381 CysGlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGly 400
DB 1340 TGCAGGCTGACAGTGTGGGCTCTGATGTACCAATCTGACAGTGCATGTGTGGG 1399
QY 401 IleValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysVal 420
DB 1400 ATCGTTAGCTGGGCTATAGTGTGGGGGCCGAGACCCGAGAGATATACCAAGGTC 1459
QY 421 SerAlaTyrLeuAsnTrpIleTyrAsnValTrpLysValGluLeu 435
DB 1460 TCAGCTTATCTCAACTGGATCTACAAATGTCTGAAAGGCTGAGCTG 1504
RESULT 4
US-10-295-027-133
Sequence 133, Application US/10295027
Publication No. US20030232350A1
GENERAL INFORMATION:
APPLICANT: Afar, Daniel
APPLICANT: Aziz, Natasha
APPLICANT: Ginsberg, Wendy M.
APPLICANT: Gish, Kurt C.
APPLICANT: Glynn, Richard
APPLICANT: Heyez, Peter A.
APPLICANT: Mack, David H.
APPLICANT: Murray, Richard
APPLICANT: Watson, Susan R.
APPLICANT: Eos Biotechnology, Inc.
TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
FILE REFERENCE: 018501-012500US
CURRENT APPLICATION NUMBER: US/10/295,027
CURRENT FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: US 09/663,733


```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/008,271
FILING DATE: 16-Jan-1998
ATTORNEY/AGENT INFORMATION:
NAME: Mohan-Peterson, Sheila
REGISTRATION NUMBER: 41,201
REFERENCE/DOCKET NUMBER: PF-0458 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-845-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 2038 Base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: COLNOT13
CLONE: 1337018
SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-10-180-719-18
Alignment Scores:
Pred. No.: 1,11e-282 Length: 2038
Score: 2338.00 Matches: 434
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.77% Mismatches: 0
Query Match: 99.83% Indels: 0
Gaps: 0
US-10-803-530-2 (1-435) x US-10-180-719-18 (1-2038)
QY 1 MetAspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysPro 20
Db 200 ATGATATCTGACATGATCAACCTCTGAACAGCTTCAGATGCAAACTCGCCCAACCC 259
QY 21 ArgLLeuProMetGlnThrPheArgLysValGlyLLeuProLLeuLLeuLeuSer 40
Db 260 CGTATCCCATGAGAGCTTCAGAAAGGTGGGGATCCCATCATCATAGCATACACTGACG 319
QY 41 LeuAlaSerLLeuLLeuValValLeuLLeuLysValLLeuAspLysTyrTyrPhe 60
Db 320 CTGGCGATATCATATTGATGTTCTCATCAAGGTGATTCGATTAATACTACTTC 379
QY 61 LeuCysGlyGlnProLeuHsIspHeLLeuProArgLysGlnLeuCysAspGlyGluLeuAsp 80
Db 380 CTTCGGGGAGAGCTCTCCACTTCATCCAGAGAGAGCTGTGTACAGAGAGCTGAGAC 439
QY 81 CysProLeuGlyGluAspGlnGluHsIspCysValLysSerPheProGluGlyProAlaVal 100
Db 440 TGTCCCTTGGGGAGAGCGAGAGCACTGTGTCAAGAGCTTCCCAAGGGCTTGCAATG 499
QY 101 AlaValArgLeuSerLysAspAArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsn 120
Db 500 GCAGTCCGCTCTCCAGAGAGCCGATCCACACTGCAAGTGTGACTCGGCACAGGAGAC 559
QY 121 TrpPheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGlnThrAlaCysArgGln 140
Db 560 TGGTTCCTGCTGCTTCGACCACTTCACAGAGCTTCGCTGAGACACCTGTAGGCGAG 619
QY 141 MetGlyTyrSerSerLysProThrPheArgAlaValGlyLLeuGlyProAspGlnAspLeu 160
Db 620 ATGGGCTACAGAGCAAACTTCATCAAGCTGTGTGAATGTGGCCCAAGCAAGATCTG 679
QY 161 AspValValGluLLeuThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyPro 180
Db 680 GATGTTGTGAATCAAGAAACACAGAGAGCTTCGATGCGAAGCTCAAGTGGGCC 739
QY 181 CysLeuSerGlySerLeuValSerLeuHsIspCysLeuAlaCysGlyLysSerLeuLysThr 200
Db 740 TGTCTTCAGAGCTCTGCTGCTCCCTGCACTGCTTGGGAGAGAGCTGAAGACC 799
QY 201 ProArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIle 220
```

```

Db 800 CCCCCTGTGTGTGGTGGGAGAGAGCTCTGTGATTTCTTGAGGAGGTCAAGCATC 859
QY 221 GlnTyrAspLysGlnHsIspValCysGlyGlySerLLeuAspProHsIspValLeuThr 240
Db 860 CAGTACGACAAACAGACAGCTGTGTGAGAGAGCATCTTGAGCCCACTGGGTCTTCAGG 919
QY 241 AlaAlaHsIspCysPheArgLysHsIspThrAspValPheAsnTrpLysValArgAlaGlySer 260
Db 920 GCAGCCACTGCTTCAGAAACATACCGATGTGTTCACTGGAAGGTGCGGCAGAGCTCA 979
QY 261 AspLysLeuGlySerPheProSerLeuAlaValAlaLysLLeuLLeuLLeuLLeuHsIsp 280
Db 980 GACAACTGGGAGAGCTTCCATCCCTGCTGTGGCCCAAGATATCATATGAAATTCAC 1039
QY 281 PrometTyrProLysAspAsnAspLLeuAlaMetLysLeuGlnPheProLeuThrPhe 300
Db 1040 CCCATGTACCCCAAGACATGACATCGCCCTCATGAAAGCTGCACTTCCCATCTTTC 1099
QY 301 SerGlyThrValArgProLLeuCysLeuProPhePheAspGluGluLeuThrProAlaThr 320
Db 1100 TCAGGCACAGTACAGGCCCATCTGTCTGCCCTTCTTGATGAGAGACTCATCAGCCAC 1159
QY 321 ProLeuTrpLLeuLLeuGlyTyrGlyPheThrLysGlnAsnGlyGlyLysMetSerAspLLe 340
Db 1160 CCACTCTGATCATTTGATGGGCTTTACGAAAGCAATGAGAGGAAATGATGTGACATA 1219
QY 341 LeuLeuGlnAlaSerValGlnValLLeuAspSerThrArgCysAsnAlaAspAlaTyr 360
Db 1220 CTGCTGACAGGCGTCAAGTCAAGTCAATTCAGACACACAGCTCATTCAGAGATCCGTAC 1279
QY 361 GlnGlyGlyValThrGluLysMetCysAlaGlyLLeuProGluGlyGlyValAspThr 380
Db 1280 CAGGGGAGATCACCGAAGAGATGATGTGTACAGGATCCCGAAGGGGGTGTGACACC 1339
QY 381 CysGlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHsIspValValGly 400
Db 1340 TCCAGGGTGCAGTGTGGGCCCTCTGATGACCAATCTGACAGTGCATGTGTGGTGGC 1399
QY 401 LLeuAlaSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrVal 420
Db 1400 ATCGTTACTGGGGTATAGCTGCGGGGCCCGACACCCAGAGATATACCAAGATC 1459
QY 421 SerAlaTyrLeuAsnTrpLLeuTyrAsnValTrpLysAlaGluLeu 435
Db 1460 TCAGCCTATCTCAACTGATCTACAAATGTCTGAAGGCTGAGCTG 1504
RESULT 3
US-11-045-577-18
Sequence 18, Application US/11045577
Publication No. US20050181404A1
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
Hillman, Jennifer L.
Yue, Henry
Guegler, Karl J.
Corley, Neil C.
Tang, Tom Y.
Shah, Purvi
TITLE OF INVENTION: HUMAN PROTEASE MOLECULES
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESS: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
```


ATTORNEY/AGENT INFORMATION:
 NAME: Mohan-Peterson, Sheela
 REGISTRATION NUMBER: 41,201
 REFERENCE/DOCKET NUMBER: PF-0458 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 650-855-0555
 TELEFAX: 650-845-4166
 INFORMATION FOR SEQ ID NO: 18:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2038 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 LIBRARY: COLANOT13
 CLONE: 137018
 SEQUENCE DESCRIPTION: SEQ ID NO: 18:
 US-09-968-415-18

Alignment Scores:
 Pred. No.: 1,11e-282 Length: 2038
 Score: 2338.00 Matches: 434
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.77% Mismatches: 0
 Query Match: 99.83% Indels: 0
 DB: 3 Gaps: 0

US-10-803-530-2 (1-435) x US-09-968-415-18 (1-2038)

Qy 1 MetAspProAspSerAspGlnProLeuAsnSerLeuAspValIleLeuAspProLeuArgIleThr 20
 Db 200 ATGATCTCTGACAGATCACTCTGAAACGCTCGATGTCMAACCCCTGGGCAACCC 259
 Qy 21 ArgIlePrometGluThrPheArgIleValGlyIleProIleIleIleAlaLeuLeuSer 40
 Db 260 CGTATCCCAAGAGACCTTCAGAAAGGTGGGATCCCATCATCATGACACTAGAC 319
 Qy 41 LeuAlaSerIleIleIleValIleValIleLeuIleValIleLeuAspIleTyrTrpPhe 60
 Db 320 CTGGGAGATCATCATGTGTGTCTCTCATCAAGTGTGATTCGATTAATACTACTTC 379
 Qy 61 LeuCySGIyGlnProLeuHisPheIleProArgIleGlnLeuCyAspGlyGluLeuAsp 80
 Db 380 CTTCGGGGCAGCCCTCCACTTCATCCCGAAGAGCAGCTGTGTGACGAGAGCTGAC 439
 Qy 81 CysProLeuGlyGluAspGluGlnIleCysValIleYSerSerPheProGluGlyProAlaVal 100
 Db 440 TGTCCCTTGGGGAGAGAGAGAGCACTGTGTCAAGAGCTTCCCGAAGGCTGACAGTG 499
 Qy 101 AlaValArgLeuSerIleAspArgSerThrLeuGlnValIleuAspSerAlaThrGlyAsn 120
 Db 500 GCAAGTCCCTCTTCCAAAGAGCCGATCCACATCGAGGTCTGGAATCCGGCCACAGGAAAC 559
 Qy 121 TrpPheSerAlaCysPheAspAsnSerThrGlnAlaLeuAlaGluThrAlaCysArgGln 140
 Db 560 TGGTCTCTGCTGCTTTCGACAACTTCACAGAAAGCTCTCGTGAGAGACCGCTGAGGACG 619
 Qy 141 MetGlyTyrSerSerIleProThrPheArgAlaValAlaGluIleGlyProAspGlnAspLeu 160
 Db 620 ATGGCTCTACAGCAGAAACCCCACTTCAGAGCTGTGAGATTGGCCAGACACAGATCTG 679
 Qy 161 AspValIleGluIleThrGluAsnSerGlnIleuArgMetArgAsnSerSerGlyPro 180
 Db 680 GATGTTGTTGAATACAGAAACAGCCAGGAGCTTCGATGCGGAATCTCAAGTGGGCC 739
 Qy 181 CysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyIleYSerLeuIleThr 200
 Db 740 TGTCTCTAGAGCTCCCTGTCTCTCTGACACTCTTTCCTGTGGGAGAGAGCTGAAGACC 799
 Qy 201 ProArgValValGlyGlyGluGlnAlaSerValAspSerTrpProTrpGlnValSerIle 220
 Db 800 CCCCTGTGTGGTGGGGAGAGAGGCTCTGTGTGATTCTTGGCTTGTGGAGGTGACATC 859

Qy 221 GlnTyrAspIySGlnHisValCysGlyGlySerIleLeuAspProHisTrpValIleThr 240
 Db 860 CAGTACACAAACAGCAGCTGTGTGAGGAGATCTCGAACCCCACTGGGTCTTCAAG 919
 Qy 241 AlaAlaHisCysPheArgIleHisThrAspValPheAsnTrpIleValAlaGlySer 260
 Db 920 GCAGCCCACTGCTTCAGAGAAACATACGATGTGTTCAACTGGAAGGTGGCCGAGCTCA 979
 Qy 261 AspIleGlnGlySerPheProSerLeuAlaValAlaIleIleIleIleGluPheAsn 280
 Db 980 GACAACTGGGACCTTCCATCTCTGCTGTGGCCAAAGATCATCATTAATTAAC 1039
 Qy 281 ProMetTyrProIyAspAsnAspIleAlaLeuMetIleLeuGlnPheProLeuThrPhe 300
 Db 1040 CCATGTACCCCAAGACATGATGATGCTCCCTCATGAAGCTGACGATCCCATCACTTC 1099
 Qy 301 SerGlyThrValArgProIleCysLeuProPhePheAspGlnGluLeuThrProAlaThr 320
 Db 1100 TCAGGACAGTCAAGCCCATCTGTCTGCTCTTTGATGAGAGCTCATCCAGCCACC 1159
 Qy 321 ProLeuTrpIleIleGlyTrpGlyPheThrIleGlnAsnGlyGlyIleMetSerAspIle 340
 Db 1160 CCACTCTGATCATTTGATGGGCTTTACGAAGCAGATGAGAGGAAATGTTGACACTA 1219
 Qy 341 LeuLeuGlnIleSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyr 360
 Db 1220 CTGCTGACAGGCGTACGTACAGGTATTCAGACGACACAGGTGCAATGCAGAGATGCGTAC 1279
 Qy 361 GlnGlyGluValThrGluIleYMetMetCysAlaGlyIleProGluGlyGlyValAspThr 380
 Db 1280 CAGGGGAGAGTCAACGAGATATGTGTGACAGCATCCCGAAGGGGCTGTGACACC 1339
 Qy 381 CysGlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValIleGly 400
 Db 1340 TGCAGGAGTACAGTGTGGGCCCCCTGATGACCAATCTGACACAGTGCAGTGTGGGC 1399
 Qy 401 IleValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValIleThrVal 420
 Db 1400 ATCGTTAGCTGGGCTATGTGTGGGGGCCCGGAGCACCCGAGAGATATACCAAGGTC 1459
 Qy 421 SerAlaTyrLeuAsnTrpIleTyrAsnValTrpIleValIleGluLeu 435
 Db 1460 TCAGCTATCTCAACTGATTCACATGTCTGGAAGGCTGAGCTG 1504

RESULT 2
 US-10-180-719-18
 ; Sequence 18, Application US/10180719
 ; Publication No. US20030166246x1
 ; GENERAL INFORMATION:
 APPLICANT: Bandman, Olga
 Hillman, Jennifer L.
 Yue, Henry
 Quegler, Karl J.
 Corley, Neil C.
 Tang, Tom Y.
 Shah, Purvi
 TITLE OF INVENTION: HUMAN PROTEASE MOLECULES
 NUMBER OF SEQUENCES: 24
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Incyte Pharmaceuticals, Inc.
 STREET: 3174 Porter Dr.
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94304
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/180,719
 FILING DATE: 25-Jun-2002

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 6, 2005, 21:57:09 ; Search time 903 Seconds
(without alignments)
3983.587 Million cell updates/sec

Title: US-10-803-530-2
Perfect score: 2342
Sequence: 1 MDPDSQPLNSLDVKKPLRKP.....VYTKVSAVLMYVWKAEL 435

Scoring table:
BLOSUM62
Xgapop 10.0 , Ygapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9793542 seqs, 4134689005 residues
Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgnt1/USPTO.spool/p/US10803530/runat_05122005_083217_2238/app_query.fasta_1.583
-DB=Published Applications NA.Main -QPMT=fastcap -SUFFIX=rmpbm -MINMATCH=0.1
-LOOPCU=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdd -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10803530 @CGN 1.1 1549 @runat_05122005_083217_2238
-NCPU=6 -ICPU=3 -NO MAP -LARGESUBV -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN_TIMEOUT=30 -THRAD=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA.Main.*
1: /cgnt2_6/prodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgnt2_6/prodata/1/pubpna/US08_PUBCOMB.seq.*
3: /cgnt2_6/prodata/1/pubpna/US09A_PUBCOMB.seq.*
4: /cgnt2_6/prodata/1/pubpna/US09B_PUBCOMB.seq.*
5: /cgnt2_6/prodata/1/pubpna/US10A_PUBCOMB.seq.*
6: /cgnt2_6/prodata/1/pubpna/US10B_PUBCOMB.seq.*
7: /cgnt2_6/prodata/1/pubpna/US10C_PUBCOMB.seq.*
8: /cgnt2_6/prodata/1/pubpna/US10D_PUBCOMB.seq.*
9: /cgnt2_6/prodata/1/pubpna/US10E_PUBCOMB.seq.*
10: /cgnt2_6/prodata/1/pubpna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2338	99.8	2038	3	US-09-968-415-18 Sequence 18, Appl
2	2338	99.8	2038	6	US-10-180-719-18 Sequence 18, Appl
3	2338	99.8	2038	10	US-11-045-577-18 Sequence 18, Appl
4	2337	99.8	1314	6	US-10-295-027-133 Sequence 133, App
5	2337	99.8	1314	6	US-10-295-027-778 Sequence 778, App
6	2337	99.8	1314	6	US-10-295-027-790 Sequence 790, App
7	2337	99.8	1314	6	US-10-295-027-830 Sequence 830, App
8	2337	99.8	1314	6	US-10-295-027-979 Sequence 979, App

9	2337	99.8	1314	6	US-10-173-999-88 Sequence 88, Appl
10	2337	99.8	2081	3	US-09-851-588-7 Sequence 7, Appl
11	2337	99.8	2104	9	US-10-956-157-2292 Sequence 2292, Ap
12	2337	99.8	2104	9	US-10-994-117-2 Sequence 2, Appl
13	2337	99.8	2104	9	US-10-991-287-2 Sequence 2, Appl
14	2337	99.8	2307	5	US-10-097-340-317 Sequence 317, App
15	2337	99.8	2307	5	US-10-171-311-217 Sequence 217, App
16	2337	99.8	2307	10	US-11-050-926-317 Sequence 317, App
17	2335	99.7	2121	7	US-10-803-530-1 Sequence 1, Appl
18	2333	99.6	2590	8	US-10-417-375-139 Sequence 139, App
19	2329	99.4	2165	6	US-10-101-510-634 Sequence 634, App
20	2328	99.4	2627	8	US-10-417-375-141 Sequence 141, App
21	2324	99.2	2070	10	US-11-071-974-2 Sequence 2, Appl
22	2324	99.2	2070	10	US-11-072-918-2 Sequence 2, Appl
23	2324	99.2	2079	3	US-09-851-588-5 Sequence 5, Appl
24	2324	99.2	2079	3	US-09-776-191-71 Sequence 71, Appl
25	2324	99.2	2079	5	US-10-264-820-22 Sequence 22, Appl
26	2324	99.2	2079	5	US-10-254-289-1 Sequence 1, Appl
27	2324	99.2	2079	6	US-10-156-214A-38 Sequence 38, Appl
28	2319	99.0	2137	3	US-09-776-191-3 Sequence 3, Appl
29	2319	99.0	2137	6	US-10-156-214A-3 Sequence 3, Appl
30	2297.5	98.1	2063	3	US-09-888-257A-2 Sequence 2, Appl
31	2297.5	98.1	2063	3	US-09-946-374-214 Sequence 274, App
32	2297.5	98.1	2063	5	US-10-006-867-111 Sequence 111, App
33	2297.5	98.1	2063	5	US-10-052-586-329 Sequence 329, App
34	2297.5	98.1	2063	5	US-10-063-547-111 Sequence 111, App
35	2297.5	98.1	2063	5	US-10-063-547-111 Sequence 111, App
36	2297.5	98.1	2063	5	US-10-176-559-329 Sequence 329, App
37	2297.5	98.1	2063	5	US-10-176-749-329 Sequence 329, App
38	2297.5	98.1	2063	5	US-10-176-758-329 Sequence 329, App
39	2297.5	98.1	2063	5	US-10-175-717-329 Sequence 329, App
40	2297.5	98.1	2063	5	US-10-063-616-111 Sequence 111, App
41	2297.5	98.1	2063	5	US-10-174-581-329 Sequence 329, App
42	2297.5	98.1	2063	5	US-10-176-481-329 Sequence 329, App
43	2297.5	98.1	2063	5	US-10-176-749-329 Sequence 329, App
44	2297.5	98.1	2063	5	US-10-176-914-329 Sequence 329, App
45	2297.5	98.1	2063	5	US-10-176-915-329 Sequence 329, App
					US-10-063-569-111 Sequence 111, App

ALIGNMENTS

RESULT 1
US-09-968-415-18
; Sequence 18, Application US/09968415
; Publication No. US20020086334A1
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; Hillman, Jennifer L.
; Yue, Henry
; Guegler, Karl J.
; Corley, Neil C.
; Tang, Tom Y.
; Shah, Purvi
; TITLE OF INVENTION: HUMAN PROTEASE MOLECULES
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/968, 415
; FILING DATE: 26-Sep-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/659, 151
; FILING DATE: <Unknown>

THIS PAGE BLANK (USPTO)

```

Db      1183 TTCTGTGCTGGCTTCATGAGAGAGTAGAGATTGATGTCAGGAGATAGTGGGGAGCCC 1242
Qy      389 LeuMetTyrGlnSerAsp---GlnTrpHisValValGlyTleValSerTyrGly 407
      1243 CATGTTACTGAAAGTGAAGGAGCAGTTCTTAACTGGAATTAATAGCTGGGGTGAAGAG 1302
Qy      408 CysGlyGlyProSerThrProGlyValTyrThrIysValSerAlaTyrLeuMetTrpIle 427
      1303 TGTGCAATGAAGGCAATATGGAATATACCAAGTATCCGGTATGTGCAACTGATT 1362

```

Search completed: December 7, 2005, 01:18:48
 Job time : 256 secs